(6)

ETL-0383

Defense Mapping Agency (DMA) raster-to-vector analysis

David D. Selden
Burton H. Went, Jr.
Stan E. Kleszczelski

BATTELLE
Columbus Laboratories
Tactical Technology Center
505 King Avenue
Columbus, Ohio 43201-2693

November 1984

Prepared for

U.S. ARMY CORPS OF ENGINEERS
ENGINEER TOPOGRAPHIC LABORATORIES
FORT BELVOIR, VIRGINIA 22060-5546

Monitored by

U.S. ARMY MISSILE COMMAND
REDSTONE ARSENAL, ALABAMA 35898-5280

85 04 29 147

E T L

Destroy this report when no longer needed.
Do not return it to the originator.

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

The citation in this report of trade names of commercially available products does not constitute official endorsement or approval of the use of such products.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS --
2a. SECURITY CLASSIFICATION AUTHORITY NA	3. DISTRIBUTION/AVAILABILITY OF REPORT Distribution Unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE NA	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) ETL-0383	5. MONITORING ORGANIZATION REPORT NUMBER(S) ETL 4
6a. NAME OF PERFORMING ORGANIZATION Battelle-Columbus Laboratories Tactical Technology Center 6b. OFFICE SYMBOL (If applicable) --	7a. NAME OF MONITORING ORGANIZATION U. S. Army Missile Command
6c. ADDRESS (City, State and ZIP Code) Columbus, Ohio 43201-2693	7b. ADDRESS (City, State and ZIP Code) Redstone Arsenal, Alabama 35898-5280
8a. NAME OF FUNDING/SPONSORING ORGANIZATION U. S. Army Engineer Topographic Laboratories Mapping Developments Division 8b. OFFICE SYMBOL (If applicable) ETL-TD-MA	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER Contract No. DAAH01-83-D-A008, D.O. No. 0022 Mods 01 and 02, MIPR No. 3.13179
8c. ADDRESS (City, State and ZIP Code) Fort Belvoir, Virginia 22060	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT NO
11. TITLE (Include Security Classification) DEFENSE MAPPING AGNECY (DMA) RASTER-TO-VECTOR ANALYSIS	
12. PERSONAL AUTHOR(S) Selden, David D., Went, Burton H., Jr., and Kleszczelski, Stan E.	
13a. TYPE OF REPORT Final 13b. TIME COVERED FROM 9/28/83 TO 11/30/84	14. DATE OF REPORT (Yr., Mo., Day) 1984, November 30 15. PAGE COUNT 128
16. SUPPLEMENTARY NOTATION	

17. COSATI CODES

FIELD	GROUP	SUB. GR.
8	2	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

Mapping, Analog-to-Vector Conversion, Cartography, Raster-to-Vector Conversion, Plotters, Digital Systems, Graphics, Benchmark Testing

Defense Mapping Agency

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

In this study a comprehensive evaluation was conducted of the analog map to digital cartographic vector data (A/V) conversion process. The evaluation consists of six major subsections: (1) development of an analog-to-vector conversion model which defines an objective framework for evaluation of systems which implement this process; (2) a study of raster-to-vector conversion algorithms and software implementations; (3) a study of current and projected A/V conversion procedures requirements at DMA Hydrographic/Topographic and Aerospace Centers; (4) a study of state-of-the-art automated cartographic data capture systems; (5) development of a standard DMA A/V benchmark testing package and methodology; and (6) recommendations for future research and development.

Additional keywords: plotters; graphics; digital systems,

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS ☒	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL MR. DOUGLAS R. CALDWELL	22b. TELEPHONE NUMBER (Include Area Code) 22c. OFFICE SYMBOL ETL-TD-MA

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

Blank Page

EXECUTIVE SUMMARY

The Battelle-Columbus Laboratories (BCL) has completed a comprehensive evaluation for the Defense Mapping Agency (DMA) of the analog map to digital cartographic vector data (A/V) conversion process. The evaluation consists of six major subsections:

1) Development of an analog-to-vector conversion model which defines an objective framework for evaluation of systems which implement this process

   o The A/V conversion model's basic components are: Preparation, Digitization, Raster-to-Vector Conversion, Feature Tagging, Spatial Coding and Data Management. Review and Edit functions are associated with each component.

   o The underlying basis for the model is a requirement that information on analog documents be converted to centerline, vector, feature tagged, spatially encoded cartographic data according to DMA standards.

2) A study of raster-to-vector conversion algorithms and software implementations

   o A discussion of pixel processing raster-to-vector conversion is presented including segmentation of raster data files into patches or swaths, line thinning, vectorization and topology reconstruction. Non-pixel processing approaches to raster-to-vector conversion is briefly discussed.

   o Alternative processing architectures are discussed including software implementation, microprogrammed implementation and hardware implementation.

3) A study of current and projected A/V conversion procedures/requirements at DMA Hydrographic/Topographic and Aerospace Centers. Pertinent results of this study include:

   o A focus on conversion requirements for unsymbolized elementary analog cartographic geometries (e.g., contours, drains, closed polygons, lines and simple "dot" point symbols), thus emphasizing cartographic features of limited symbolization or geometric complexity.

   o A highlighting of typical geometric degradations, errors and anomalies (e.g., gaps, spikes, snow, centerline misalignment) either encountered on input manuscripts or resulting from raster scanning/raster-to-vector conversion.

   o A reaffirmation of the need to evaluate all components of the cartographic data capture process (as defined in the A/V model) to pinpoint critical deficiencies and understand trade-offs between alternative solutions. This may be exemplified by the allocation of human resources to intensive material preparation (particularly for removal of errors and anomalies) as an alternative to development of sophisticated software for automatic error detection/correction.

4) A study of state-of-the-art automated cartographic data capture systems. Results of this study include:

   o The development of a standard questionnaire for profiling commercial/research A/V and R/V systems according to the major components of the A/V conversion model.

   o The development of standardized profiles for all major "cartographic" raster scanning and automatic line-following systems currently in the marketplace or being utilized by industry.

- o A better understanding of some of the critical differences between systems and how they might best serve DMA requirements.

5) Development of a standard DMA A/V benchmark testing package and methodology

   - o The material testing package includes representative DMA analog manuscripts including DTED contour/drain-ridge overlays, a DFAD color pencil Mylar compilation, and a color pencil Mylar hydrographic chart compilation.

   - o The material testing package also includes two types of synthetically generated analog inputs: a) twelve basic cartographic geometries (non-intersecting lines, intersecting/merging lines, and intersecting/crossing lines) each plotted in four increasing levels of density and b) one sheet representing samples of "perfect" cartographic geometries with examples of geometric degradation along side.

   - o The benchmark testing methodology consists basically of scanning (or digitizing) the twelve synthetic sheets described above and vectorizing them while keeping timing statistics. This provides information about a systems level of performance in processing different geometries and the impact of increasing data density.

   - o The second part of the benchmark testing consists of scanning and vectorizing the synthetic test sheet in category (b), described above. In addition to timing these processes, the running, timing and quality assessment of automatic data editing procedures is required.

6) Recommendations for future research development

- The development of an advanced benchmark testing capability is recommended. This would be based on the benchmark package resulting from the current research effort. It would address technological advances in the areas of pattern/symbol/character recognition, automatic feature tagging/contour elevating and computer-assisted spatial coding procedures.

- A feasibility study is recommended for the development of an optical image processing capability for cartographic data capture. The development of an intelligent scanner for automatic feature recognition and feature tagging utilizing optical image correlation technology deserves further attention.

- A program of vendor system optimization and upgrade is recommended. A systematic review of existing cartographic data capture systems at DMA (e.g., Scitex) with a goal of more fully and effectively utilizing available functions represents a significant contribution to enhanced productivity.

- Additional research and development programs are recommended in the areas of advanced computer architectures for cartographic data processing, raster-to-vector conversion algorithms and cartographic data structures.

PREFACE

This research was supported by the U. S. Army Engineer Topographic Laboratories, Mapping Developments Division, Fort Belvoir, Virginia, and was monitored by the U. S. Army Missile Command, Redstone Arsenal, Alabama, under Contract No. DAAH01-83-D-A008, which is sponsored by the Defense Advanced Research Projects Agency.

Battelle is not engaged in research for advertising, sales promotion, or publicity purposes, and this report may not be reproduced in full or in part for such purposes.

The use of trade names in this report does not constitute an official endorsement or approval of the use of such commercial hardware or software.

ACKNOWLEDGMENTS

The authors wish to thank the following for their assistance in preparing this report: Christina Weber, Terri Stiteler, and Audrey Buggs for secretarial activities, and Ronald Exler for editing and graphics assistance. Special thanks are offered to the U. S. Army Engineer Topographic Laboratories; Defense Mapping Agency (DMA), Hydrographic/Topographic Center; DMA, Aerospace Center; and DMA, Headquarters, for their valuable suggestions and critical reviews of the effort.

<u>Note to the Reader</u>: **Chapters 4.0 and 5.0 of the report are included in an appendix issued under separate cover. This appendix, entitled "Special Appendix to the Defense Mapping Agency (DMA) Raster-to-Vector Analysis", is limited to distribution within the Department of Defense. The page numbering scheme for the main report and the special appendix has been maintained in sequential order to enhance the readability of the report in its entirety. Pages 51 through 74 comprise Chapter 4.0 and are located in the special appendix. Pages 75 through 78 comprise Chapter 5.0 and are also located in the special appendix.**

DISCLAIMER

Blank Page

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

APPENDICES

FIGURES AND TABLES

1.0 INTRODUCTION

Battelle-Columbus Laboratories (BCL) has completed a comprehensive evaluation for the Defense Mapping Agency (DMA) under Contract No. DAAH01-83-A008 entitled "A Defense Mapping Agency (DMA) Raster-To-Vector Analysis". The following technical report summarizes the results of our investigation and lays the groundwork for further research and development.

1.1 Background

The Defense Mapping Agency has a significant requirement to convert large quantities of analog cartographic manuscripts to various kinds of digital computer representation. This work is currently performed by the Defense Mapping Agency Hydrographic/Topographic Center (DMAHTC) in Washington, D.C. and the Aerospace Center (DMAAC) in St. Louis, Missouri. Both production centers employ a mix of commercially available technologies to complete this map conversion task. Present state-of-the art methodology is implemented by raster scanning of cartographic manuscript overlays, raster-to-vector conversion and feature tagging as a post-processing procedure.

1.2 Problem Definition

The current DMA cartographic data production process encounters significant bottlenecks. Raster scanning of analog cartographic manuscripts containing large numbers of point, line and area features (unsymbolized) produces a dramatic increase in throughput rates for initial data capture, in comparison to manual digitizing techniques. However, procedures for source manuscript preparation, raster-to-vector conversion, error detection/ correction, and feature tagging are particularly time consuming and error-prone.

1.2.1 Statement of Work (SOW) Requirements

The Statement of Work for the Defense Mapping Agency (DMA) Raster-To-Vector Analysis project defines three primary objectives; 1) Task #1: An Analysis of DMA Raster-to-Vector System Requirements, 2) Task #2: A State-of-the Art Raster-to-Vector Analysis, and 3) Task #3: The Development of Standard Definitions and Testing Methods.

1.2.2 Expanded Problem Focus

Battelle proposed at the Raster-To-Vector Analysis "kickoff" meeting[1] an expanded problem focus and a reordering of task priorities. An outline of a conceptual approach to evaluating the raster-to-vector process was presented which included proposals for the following: a) The definition of an Analog-to-Vector Conversion (A/V) Model with standardized terminology to describe the process, b) The description and evaluation of current DMA analog-to-vector conversion processing in terms of the A/V model, c) The description and evaluation of the future DMA A/V requirements, d) The description, evaluation and analysis of state-of-the-art analog-to-vector and raster-to-vector conversion systems in terms of the A/V model, and e) The development of a benchmark testing package and standard methodology for evaluation of state-of-the-art A/V and R/V systems.

1.2.2.1 Historical Perspective

Recent history at the Defense Mapping Agency has witnessed an accelerated evolution of digital cartographic production techniques, much of which has been focused on the production of map/data products derived from

[1] November 7, 1983 at DMAHTC.

original or photogrammetrically compiled analog cartographic manuscripts. The generation of elevation matrix data sets (DTED cells) for example, originally required the tracing of contour lines with the Digital Graphics Recorder (DGR) stylus "cursors". Elevating and feature tagging were performed concurrently with digitization and the resulting output was centerline digital cartographic data. This elegant and simple process was deemed too slow to meet DMA's production requirements however, and alternative methods were sought.

The Broomall Scan Graphics Automated Graphic Digitizing System (AGDS) was acquired to replace the DGR "system" with expectations for significant throughput improvement with a reduced manpower requirement. The AGDS employs raster scanning and processing technology to convert cartographic source materials to vector cartographic computer data. This task is accomplished through a configuration of hardware and software built into three subsystems; raster scanner, vectorizer, and edit/tag. Essentially, this system is designed to perform tasks identical to those performed by its predecessor, only faster.

Although no hard data is available, it is widely assumed that the AGDS generates digital cartographic data (which meets DMA's quality standards) at a higher production level than the DGR system. These production levels can be further improved, thus the need for further research and development. As mentioned above, raster scanning does produce large amounts of digital cartographic data in relatively short time frames. It is the ensuing data processing which is time consuming, error prone and entirely too labor intensive. DMA has dedicated considerable resources during the past years to address the deficiencies of these data processing methods.

The more recent acquisition of the Scitex Response-250 represents a significant action in this regard. Considered by many to support state-of-the-art hardware and software dedicated to the conversion and processing of analog cartographic materials, this system provides DMA with greatly enhanced capabilities. Currently, the Scitex is utilized for color separation scanning, hydrographic chart compilation, revision and production, in addition to slope mapping and other topographic mapping support. Its optimal support role is often re-evaluated in light of changing DMA requirements.

The development of improved raster-to-vector algorithms and computer implementations has also been the focus of much of DMA's research and development. The Environmental Research & Technology, Inc. "Mini Raster-to-Vector Conversion Program" and the Goodyear Aerospace Corporation's "Associative Array Processing of Raster Scanned Data for Automated Cartography" are two examples of these efforts.[2]

Battelle's major concern with these and similar efforts is that their applicability to DMA production requirements is limited by their view of the total production process. They also do not directly support DMA's evaluation of in-house commercial cartographic data capture systems and other state-of-the-art systems in today's marketplace. It is with this in mind that Battelle proposed an evaluation of the raster-to-vector "problem" as part of an overall process; the analog map to digital cartographic vector data conversion process (A/V).

1.2.2.2 Analog-to-Vector Conversion

The conversion of analog cartographic point, line and area features to digital cartographic data results from a process which coordinates trained personnel, hardware and software configurations. We will refer to this process as analog-to-vector conversion (A/V). A/V consists generally of manuscript input preparation, digitization, raster-to-vector conversion, tagging, spatial coding and data management. All these components must work efficiently and according to prescribed quality standards in order to attain a high level of productivity. Little is gained by developing a procedure which exhibits a high

2 Refer to Appendix E for a description of previously sponsored DMA raster-to-vector analysis research and development.

degree of efficiency (e.g., raster scanning) while lacking the capability to process its output to an equal degree. Similarly, the development of a raster-to-vector conversion software routine which improves such conversion dramatically is equally suspect if it creates output data far below acceptable quality standards.

1.2.2.2.1 Raster-To-Vector Conversion. Raster-to-vector conversion (R/V) is a unique process, required if raster scanning has been employed for initial data capture. R/V is a sub-set of the analog-to-vector conversion process. Procedures prior to its implementation have a significant impact on its efficacy (as do data types) and its own processing will influence subsequent digital cartographic production requirements.

An approach to evaluating or developing algorithms for raster-to-vector conversion must be considered in light of the overall cartographic process and immediate expectations. For example, there is a significant difference between a straight-forward requirement to convert unsymbolized, linear cartographic data from a raster data structure to a center-line vector representation and a requirement to produce similar results with a fully symbolized cartographic raster input. The ability to automatically produce center-line vector data which is also tagged with an attribute represents another significant technological advance. Thus, it is obviously unfair to compare throughput rates of various R/V algorithms, for example, if they perform tasks at different levels of complexity. The overall consideration must be to determine what role raster-to-vector most effectively plays as part of the analog-to-vector process in producing feature tagged, spatially coded, center-line digital vector cartographic data.

2.0 ANALOG-TO-VECTOR CONVERSION: DEFINITIONS AND ISSUES

This section provides a comprehensive definition of the major components which form the analog-to-vector conversion process. Each component in this cartographic process is addressed separately, highlighting pertinent functions, parameters, and issues. In doing so we are better able to evaluate the Defense Mapping Agency's particular requirements and develop useful procedures for the description, evaluation and analysis of state-of-the-art A/V and R/V systems.

2.1 Analog-to-Vector Conversion Model

Battelle, in consultation with the Defense Mapping Agency, has developed an analog-to-vector conversion (A/V) model. The impetus for the creation of this model was the need to form a conceptual framework within which the raster-to-vector conversion problem could be more effectively evaluated. Although this model can be broadly applied to the generic digital cartographic production process, it has been developed to apply specifically to digital cartographic production systems at DMAHTC and DMAAC. (A/V Conversion Model - Figure 1A).

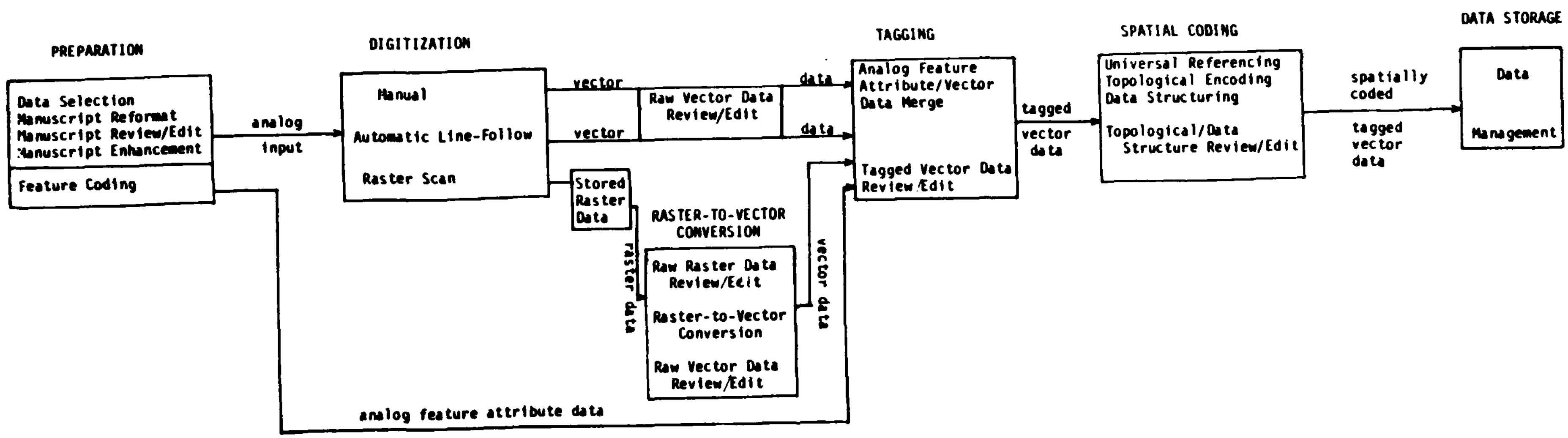
2.1.1 Analog-To-Vector Conversion Model Components

The analog-to-vector conversion model consists of five essential processes. These are: Preparation, Digitization, Tagging, Spatial Coding, and Data Management. The sixth process, Raster-to-Vector Conversion, is implemented where raster scanning is employed for initial cartographic data capture.

PREPARATION subdivides into five functions, four concerned with the analog cartographic manuscript and the other, feature attributes. Data Selection, Manuscript Reformat, Manuscript Review/Edit, and Manuscript Enhancement comprise the four manuscript-related functions of preparation. Feature Coding is the fifth preparation procedure.

Figure 1A

ANALOG-TO-VECTOR CONVERSION MODEL



DIGITIZATION is typically performed in one of three modes. These include Manual, Automatic Line-Following and Raster Scanning. Although manual digitizing continues to serve a useful function it is the latter two approaches (particularly raster scanning) which concern this evaluation.

RASTER-TO-VECTOR CONVERSION contains three functions as defined by the model. Raw Raster Data Review/Edit, Raster-to-Vector Conversion and Raw Vector Data Review/Edit are all performed under digitization.

TAGGING is dedicated to the merging of analog feature attribute information (usually in the form of numeric codes) with center-line vector cartographic data. This is descriptively referred to as Analog Feature Attribute/Vector Data Merge. Tagged Vector Data Review/Edit is another sub-function of Tagging.

SPATIAL CODING consists of three main sub-functions and a quality assurance function. The three basic functions are Universal Referencing, Topological Encoding and Data Structuring. The quality assurance function is described as Topological/Data Structure Review/Edit.

DATA MANAGEMENT is the final process and for the purposes of the analog-to-vector model it is not broken down to sub-functions.

2.1.2 The Analog-to-Vector Conversion Model Glossary

Battelle, in consultation with DMA, has developed a glossary of terms which defines the individual components of the analog-to-vector conversion model. The standardization of terminology affords Battelle (as well as DMA) the opportunity to evaluate systems which perform analog-to-vector and raster-to-vector conversion in an objective and consistent manner.

See Appendix A for the current version of the Analog-to-Vector Conversion Model Glossary.

2.2 Analog-to-Vector Conversion: Preparation

Preparation, the first component of the analog-to-vector conversion process, has been fully defined in the A/V Model Glossary (see Appendix A). The following sections discuss two aspects of this process: generic data types and generic input data/material characteristics; their impact on digitization, raster-to-vector conversion, feature tagging, and quality assurance procedures warrants further investigation.

2.2.1 Generic Data Types

Seven generic types of data have been identified, each representing unique cartographic morphologies. These include: non-intersecting lines, closed polygons, merging/intersecting networks, crossing/intersecting networks, point features, symbolized elements, and labeling. Each type presents a unique set of problems and challenges to map conversion methodologies. (See Figures 2A, 2B, and 2C for graphic examples.)

2.2.1.1 Non-Intersecting Lines

A self-enclosing line (with starting and ending points the same) of equal data values may be referred to as a non-intersecting line. Contours are the predominant example of this type of data at DMA. Other examples include isotherms (lines of equal temperature) and isohyets (lines of equal rainfall value).

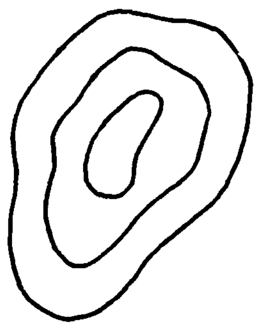
2.2.1.2 Closed Polygons

Areal cartographic features formed by merging curvilinear elements may be referred to as closed polygons. Digital Feature Analysis Data (DFAD) compilation manuscripts contain this type of data. These polygons may stand alone, merge with other polygons (but not intersect), and may be subsumed by larger polygons. Other examples of closed polygons are found on soil maps and land use/land cover maps.

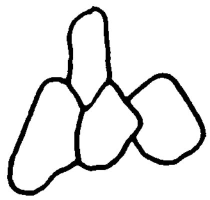
Figure 2A

Generic Data Types (Cartograpic)

Non-Intersecting Lines



Closed Polygons



Merging/Intersecting Networks

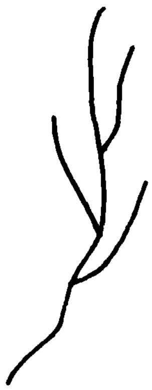
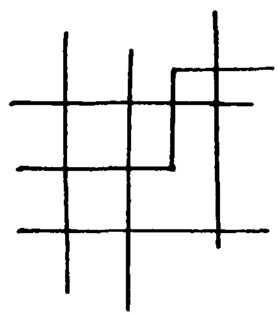
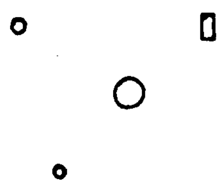


Figure 2B

Crossing Intersecting Networks



Point Features



Symbolized Features

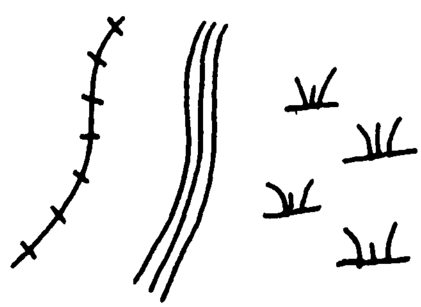
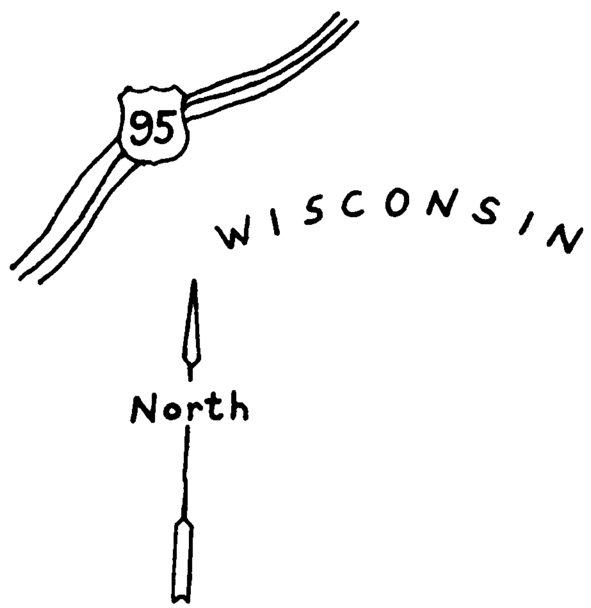


Figure 2C

Labeling



2.2.1.3 Merging/Intersecting Networks

Linear and crenulating features which merge (but do not cross) are referred to as merging/intersecting networks. Streams and drainage patterns are prime examples of merging networks. These are very evident on standard DMA topographic maps. They are also captured with contour overlays for the generation of Digital Terrain Elevation Data (DTED).

2.2.1.4 Crossing/Intersecting Networks

Linear features which merge and cross form crossing/intersecting networks. Roads are a typical form of intersecting networks. Graticules are another example. Roads can be located on many DMA map products including topographic maps and hydrographic charts.

2.2.1.5 Point Features

Features whose identification is concerned with a singular geographic location, as opposed to linear or areal cartographic features which define multiple data point locations are referred to as point features. DFAD and DVOD manuscripts contain point features which represent the geographic center of radar identifiable features on the earth's surface. Topographic and hydrographic products contain many point features, including: housing, cultural and physical landmarks, buoys and lighthouses.

2.2.1.6 Symbolized Elements

All geographic phenomenona are symbolized to some extent if they are portrayed as cartographic features on a map. Symbolized elements consist of those characteristics of a cartographic feature or symbol which facilitate its detection, recognition, differentiation, and identification. A wide range of symbolic complexity is evidenced by symbolized elements. Increased complexity often causes greater difficulties in the conversion of these elements to feature tagged, digital center-line vector data. All generic data types possess symbolized elements in varying degrees. This study addresses the problems associated with "capturing" and processing generic cartographic data types containing the lowest level of symbolized elements (i.e., point, line and area features of varying lineweights only).

2.2.1.7 Labeling

Labeling refers to all cartographic names and alphanumeric symbols portrayed on a map. Practically all map products contain such labeling as a means to enhance communication. The difficulty of assimilating this data type into the analog-to-vector conversion process cannot be overlooked. Labeling is a predominant cartographic component and plays a significant role in the building of "intelligent" digital cartographic data sets. (This study does not address the problems associated with "capture" and processing cartographic labeling.)

2.2.2 Generic Input Data/Material Characteristics

Six analog manuscript/data characteristics have been identified which influence procedures for material preparation, digitization and raster-to-vector conversion. These are material type, data representation type, material age, data density, data quality, and material-data/system(s) compatibility.

2.2.2.1 Material Type

The type of analog manuscript material may present some difficulties in the A/V process. Different types of materials commonly used are paper, Mylar, Scribecote, film positives and film negatives. Paper may be problematic due to geometric instability. Other materials may be unacceptable due to scanner incompatibility. For example, the grain of a particular brand or batch of Mylar will affect the resolution and reflectance requirements of different raster scanners.

2.2.2.2 Data Representation Type

Data representation type refers to the kinds of material used to represent cartographic features on the manuscript. Typical data representation types include: pencil ("lead"), color pencil, black ink ("wet"), four color ballpoint, photo-black, printed color ("line"), and printed color ("tint screen"). Analogous to material type, certain kinds of data representation are problematic. Lead pencil may not be resolvable by certain scanners, resulting in missing data, broken lines, and uneven lineweights. Multiple colors on a single manuscript are only appropriate for color recognition scanners. Specific limitations for the number of different colors permissible on a single overlay differ from one scanner to the next.

2.2.2.3 Material Age

If analog cartographic manuscripts are old they may begin to deteriorate and exhibit various types of problems. Long periods of storage under different conditions may result in paper shrinkage, folding, tearing and destruction of data representation images. Different types of materials are adversely affected by age to varying degrees. Paper is most drastically affected. Stable base materials such as film and Mylar are less affected.

2.2.2.4 Data Density

The density of data on a cartographic manuscript is important from two vantage points. First, if the data coverage is sparse, the choice of manual digitization may be more efficient than raster scanning. Second, if cartographic features are too close together (e.g., contours) it may pose problems for raster scanners in their ability to resolve the separation of the features. Data density may also impact on scanning time for some scanners and obviously influences raster-to-vector conversion, feature tagging and quality assurance activities.

2.2.2.5 Data Quality

The quality of input data is critical to the efficient processing of cartographic manuscripts. Increasing numbers of manuscript errors and anomalies greatly encumber all data production procedures. Errors such as gaps, spikes and miscellaneous marks directly affect raster-to-vector conversion and editing procedures. The key question is whether it is more effective to prepare (or "re-prepare") manuscript inputs according to the highest quality standards (and compatible with data capture system requirements) or submit sources in their existing condition; the latter requiring the development of sophisticated editing techniques to resolve problems later in the process. The answer lies partially with the unique requirements of a particular organization. In those cases where manpower reduction is a serious concern, manuscript re-preparation is not a viable option. This does require, however, the development of automated quality assurance and editing procedures, to reduce the transference of labor intensive input manuscript preparation to labor intensive quality assurance and editing activities.

2.2.2.6 Material-Data/System(s) Compatibility

All the above mentioned parameters combined, influence material/system compatibility. The compatibility issue must be considered on a case by case basis taking into account the basic characteristics of the input manuscripts and the inherent limitations of particular data capture systems. Some digitizing systems recognize multiple colors and others do not. Some raster scanners have variable resolution settings and others do not. Some data capture systems have advanced quality assurance and editing algorithms and others do not. One example of material/system incompatibility concerns material format size. Currently, DMA hydrographic charts require folding into sections to permit scanning on the Scitex raster scanner. (This would be the case for most other commercial scanners and certainly all color recognition scanners in today's market).

2.3 Analog-to-Vector Conversion: Digitization

Various methods of digitization are available in today's market. DMA employs a mix of these systems in its data capture operations. This section will discuss the alternative methods of digitization and describe those characteristics which define a particular system.

2.3.1 Alternative Methods

Presently, a number of methods are available which perform the conversion of analog cartographic information to some form of computer representation. Three major types of data collection techniques exist, each exhibiting variations within it. The three generic approaches to digitization include: manual, automatic line-following, and raster scanning/raster-to-vector conversion.

2.3.1.1 Manual Digitization

The "manual" process of tracing point, line and areal cartographic features has historically been the mainstay of digital cartographic data capture. Its advantages are: simplicity of technique, the ability to perform feature tagging and quality checks concurrently, and direct output of center-line vector data. Its disadvantages include: a labor intensive process, prone to errors (i.e., human operator eye/motor control limitations), and its time consuming nature (as a function of data density). It should be noted that alternative manual digitizing methods have been developed, including error correcting and line-following approaches.[3] Overall, the advent of more automated data capture systems and procedures has greatly reduced the dependence on all manual techniques. However, manual techniques continue to play effective roles under certain conditions.

The digitization of point feature locations is currently more effectively accomplished using manual digitizing techniques. Many raster scanning systems have difficulty capturing and storing individual point

---

[3] Vogel, James M. "In Laboratory Investigative Research: Intelligent Line Digitizing Study Interim Report," 1983.

features, either eventually "losing" them or storing them as short line segments. The latter results in an ambiguous definition of the feature's geometric center. The other condition where manual digitization is advantageous is when data coverage on a particular cartographic manuscript is sparse. Although no firm guidelines are available (and depending on what alternative automatic system is available), it does appear to be more efficacious to employ manual digitizing techniques below a certain threshold of cartographic elements. At DMAAC, for example, Digital Feature Analysis Data (DFAD) compilation manuscripts containing approximately one hundred and thirty[4] features or less are digitized on the AGDS edit/tag subsystem instead of being scanned on the raster scanner subsystem.

2.3.1.2 Automatic Line-Following

This technique is fundamentally an extension of manual digitization. Conceptually, the following of cartographic linear features conforms to, and reinforces the traditional "vector" view of cartography, only in a digital computer environment. The most outstanding example of the automatic line-following approach is the Laserscan Lasertrak which utilizes deflected laser beams in a local raster scanning, line-following collection of the x,y coordinates of cartographic elements. One significant advantage of this approach is the provision for concurrent feature tagging and intelligent operator interaction with a computer-assisted digitization process. Thus, the resulting data is not only in a "desirable" vector format, but feature tagged as well. Here again this technique appears to have certain limitations and should be utilized only where appropriate. The Laserscan Lasertrak, for example, is very effective in the data capture of non-intersecting lines, such as contours. It is also reasonably effective in the digitization of closed polygon maps. Dense networks (roads, for example) have been reported to cause this approach some difficulty, at least from an efficiency point of view.[5]

---

[4]According to estimate made by DMAAC staff.

[5]No quantitative information available

The inability to efficiently capture cartographically symbolized features (other than simple dashed patterns) is another limitation, although this problem is not restricted to the automatic line-following approach.

2.3.1.3 Raster Scanning/Raster-to-Vector Conversion.

Three variations of this basic approach include: raster scanning/raster-to-vector conversion (two step software), raster scanning/raster-to-vector conversion (realtime software), and raster scanning/raster-to-vector conversion (realtime "hardware"). All three have the potential to be implemented as basic ("dumb") or advanced ("intelligent") procedures. (See Sections 2.4.2.1 and 2.4.2.2 for further discussion).

2.3.1.3.1 Raster Scanning/Raster-to-Vector Conversion (two step - software).

This two step process commences with the raster scanning of an analog cartographic manuscript, resulting in a digital raster cartographic "image". This digital image is then converted from a raster, scan line structure, to a vector center-line data representation through the application of software routines, usually on a separate system. Some commercial systems provide computer-assisted and automatic software functions to edit raw raster data prior to raster-to-vector conversion. Others provide similar capabilities in vector mode after the conversion is completed. Some provide both capabilities.

2.3.1.3.2 Raster Scanning/Raster-to-Vector Conversion (near realtime - software).

This approach combines raster scanning and raster-to-vector conversion (software) in near realtime. (Vectorization starts during scanning but does not complete until after completion of scanning.) The main advantage of this approach is an expected increase in throughput. However, current implementations of this concurrent processing (e.g., Broomall Scan Graphics - RAVE) appears to preclude the automatic editing of raw raster cartographic data. This is particularly critical in light of recent evidence pointing toward the development of more efficient editing routines in raster mode.[6] If the increased throughput gained by this method results in the generation of uncorrected errors and anomalies, such gains may be neutralized or even lost due to corrective requirements later in the process.

2.3.1.3.3 Raster Scanning/Raster-to-Vector Conversion (realtime - hardware).

This represents an alternative implementation of raster-to-vector conversion. Two unique approaches have been identified. The first approach utilizes a hardware processor to convert raster scan data to boundary chains concurrent with raster scanning, followed by software generation of centerline vectors.[7] The second approach scans analog manuscripts in patches and converts each raster data patch to centerline vector data immediately, via a hardware processor.[8] Both approaches claim dramatic improvements in raster-to-vector conversion speed (although no similar claims are made for quality improvement). Of particular concern to the cartographic community, however, is the type of vector data structure which results from these processes, and its direct applicability to cartographic data base requirements. This concern is specifically directed to the first approach described above. A second issue concerns the applicability of such hardware processing to the broad range of cartographic geometries found on typical cartographic manuscripts. Questions remain as to the flexibility of such approaches to raster-to-vector conversion. Empirical testing and specification of vector data structure formats will be required to validate the usefulness of these hardware approaches.

---

[6]Note Scitex, Sysscan, as leaders in this direction.

[7]ANA Tech, Teledyne Geotronics.

[8]Teledyne Geotronics has developed a system based on this approach.

2.3.2 System Characterization

Battelle has identified seven parameters which together characterize a data capture or digitizing system. These include: hardware overview, installation requirements, digitizing method, image holder characteristics/analog input requirements, photometric characteristics, accuracy specifications, and productivity.

2.3.2.1 Hardware Overview

This is an overview of the basic system components and their specifications. Major components of typical digitizing systems include: raster scanner (drum, flatbed), digitizer (cursor, trackball, light pen), mini-computer, CRT, tape drive, disk drive and disk storage, digitizing table (pad, tablet), and hardcopy output device.

2.3.2.2 Installation Requirements

Installation requirements are concerned with five parameters: floor space, system weight, power supply, environmental requirements, and required computer support.

2.3.2.3 Digitizing Method

As previously mentioned three generic alternative digitizing methods are available in today's market: manual, automatic line-following, and raster scanning. Within these three categories we find different technological solutions. In the manual category three types currently exist: standard ("hand-driven"), line-following (where a line is manually identified and a cursor is "locked" onto it and then automatically followed to its conclusion), and error correcting ("hand-driven" with a built-in microprocessor which automatically adjusts to deviations in maintaining a center-line track within a prescribed radius). Automatic line-following digitizers appear to fall into the category of deflected laser tracking systems. Raster scanning systems typically employ three different approaches for their "cameras": electro-optical, laser, and solid state (with either flatbed or drum image carriers).

2.3.2.4 Image Carrier Characteristics/Analog Input Requirements

Three types of information are collected under this heading: image carrier, registration system, and format specifications. Image carrier typically divides between drum and flatbed. Registration system either takes the form of tabs or fiducial marks located on the input manuscript. Some scanning systems provide a hold-down glass, but this is not a precise registration system per se. There are two format specifications worthy of note: sheet format and image format. Sheet format is the maximum manuscript sheet size which can be mounted onto the image carrier, laying flat and firmly secured. Image format is the maximum dimensions of the cartographic image which can be captured in a single pass.

2.3.2.5 Photometric Characteristics

According to George Nagy (University of Nebraska) in his article "Optical Scanning Digitizers" (IEEE 1983) photometric characteristics include: range, sensitivity, uniformity, stability, repeatability, and spectral response. Range refers to the "difference in optical density corresponding to the saturation and dark thresholds of the device." Sensitivity indicates the smallest measureable difference in grey level. Uniformity is the measure in the variation of photometric response to different points on an image which produce equal output. Stability is the long term variability in photometric response to a single location on an image. The same definition applies to repeatability. Spectral response "is the amplitude response of the scanner as a function of the reflected or transmitted light." This final parameter is basically referring to a scanner's capability to sense color, grey scales or black and white images.

In terms of cartographic applications photometric characteristics are important in the following areas: color recognition, color storage, color calibration, grey level recognition, grey level storage and grey level calibration. These all vary dramatically from system to system.

2.3.2.6 Accuracy Specifications

Accuracy is important in maintaining the geometric integrity of cartographic data. Accuracy concerns two phenomena: geometric linearity and stability. George Nagy defines geometric linearity as "the deviation from the ideal that occurs when a perfect straight line is reproduced. For instance, 0.1 percent geometric linearity means that the digitized version of a 10-cm line segment can deviate from its calculated position at most by the number of pixels needed to make up 0.1 mm." Geometric stability is measured in the variability in geometric linearity over time. Short term stability is determined by making measurements on a cartographic manuscript without removing it from its holder. Long term stability is determined from test to test (over time, where a document is removed) and usually is a function of the image holder.

2.3.2.7 Productivity

Productivity is a measure of how many units of a particular product are produced in a given amount of time. Different measures may be used depending on the digitizing system employed. Automatic line-following systems, for example, can be measured for productivity by analyzing line-following rates. Scanning systems require an analysis of traverse rate, image size, spot size and sampling resolution. The impact of data density also affects certain systems. Others are affected by segment convolution, (particularly line-followers). It is important to note that the significance of speed is mitigated by quality of initial output. High speed and quality ratings are equally important factors in productivity analysis.

2.4 Analog-to-Vector Conversion: Raster-to-Vector Conversion

The A/V Model Glossary identifies three steps in raster-to-vector conversion: raw raster data review/edit, raster-to-vector conversion, raw vector data review/edit. Pertinent topics and parameters will be discussed in this section in terms of characteristics which define raster-to-vector systems and procedures.

2.4.1 Raw Raster Data Review/Edit

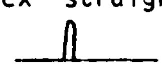
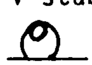
Two issues are important in evaluating this function: data storage/compaction and error detection/correction. The variation in approaches taken by different systems to performing this step is significant. Recent trends indicate that the storage and manipulation of raster scanned data will play an increasingly important role in the future.

2.4.1.1 Data Storage/Compaction

Raster scanning of cartographic source documents produces large quantities of data. Depending on the resolution, the total amount of data collected would overwhelm most processing software, if left untouched. Data compaction is critical to efficient processing and conversion. Two common approaches to data compaction are run length encoded and entry/exit point storage. Run length encoding essentially stores a counter for every continuous string of equal values, often indicated by 0 or 1 bit indicators. This avoids having to store every data location. It does require the "rebuilding" of vectors during raster-to-vector conversion, however. Entry/Exit storage stores a start scan line location and only indicates the location of a threshold change (entry), and its subsequent change (exit). Simply put, raster locations are stored only when the existence of a new cartographic element is detected. Line-following systems also compact their data before further processing. This is usually accomplished through filtering and point reduction schemes. Various types of algorithms result in the storage of key inflection points critical to the faithful representation of the cartographic features.

2.4.1.2 Error Detection/Correction

Most raster systems provide some means of reviewing the results of scanning a cartographic manuscript, if only to assure that elementary quality standards are met and that the entire image has been captured. The significant difference is in the provision of further capabilities to perform actual edits on the raster data. Three forms of editing capabilities have been identified: interactive, computer-assisted, and automatic. Interactive editing is achieved through the facilities of computer interactive graphics. Each error or anomaly is individually identified by a human operator and corrected through the use of a light pen, cursor or trackball. Computer-assisted editing occurs when a software routine identifies errors or anomalies (usually based on parameters pre-set by the human operator), brings the error in question to the attention of the operator (either bringing the element to the center of a CRT or through color coding), and the human operator then corrects the problem through the use of a light pen, cursor or trackball. Automatic editing occurs through software routines which identify and correct different types of errors found in raw raster data. Some of these routines provide for input from human operators who enter different parameters and tolerances relevant to the particular data source.

The complexities of cartographic data anomalies or errors make successful automation of error detection/correction algorithms very difficult. Observation of one type of data anomaly, stubs (also referred to as spikes) for example, elucidates this point. Eradication of the numerous permutations of data stubs (primarily caused by the thinning of variable width raster data elements) is not accomplished by simply setting a pixel length/width tolerance. This may succeed for the least geometrically complex "straight stub" , "v stubs" , "circular stubs" , "tube stubs" , "donut stubs"  and other variations. Similar complexities are encountered with the full range of error types. Although state-of-the-art data capture systems provide automated routines for addressing these problems, success rates may vary significantly. The availability of the three basic editing approaches discussed above is necessary for handling these complex challenges.

2.4.1.2.1 Error Types. Errors and anomalies evidenced in cartographic data are attributable to cartographic source materials, raster scanning, and line thinning/vectorization. Analog source documents contain features with gaps or spikes, lines of varying thickness or optical density, and dirt spots which can become "unwanted" data during subsequent processing. The resolution of a raster scanner, its grey level/color recognition capabilities or its calibration may result in "lost" data, merged data or undifferentiated data types. Line thinning of raster cartographic data often results in discontinuities (gaps), spurs (often caused by uneven lineweight input) and "unthinned" lines particularly those of strong angular trajectory. The combination of large numbers of different types of errors greatly complicates editing tasks. The charts in Appendix B portray abstractions of a representative sample of error types. The charts include a short description of each error type, a graphic simulation, a statement of possible error cause and a statement of alternative error correction approaches.

2.4.2 Raster-to-Vector Conversion

As indicated in the A/V Model Glossary this represents the actual conversion of raster formatted cartographic data to vector center-line representation. There are two approaches to this conversion process: the basic approach has been implemented in most raster systems since their inception. The advanced approach is currently under development in a number of leading commercial concerns.

2.4.2.1 Basic Conversion

This approach provides for the conversion of raster cartographic data to a digital vector center-line representation irrespective of their unique characteristics or data category. For example, if a fully symbolized railroad feature is raster scanned and passed through a basic R/V routine, the resulting data will maintain center-line vectors for each individual railroad tick. For most cartographic data bases this represents an enormous data storage burden, and is impractical from a topological point of view as well. These vector data also remain "dumb" due to a lack of feature identification.

2.4.2.2 Advanced Conversion

Advanced raster-to-vector conversion provides for the generation of center-line vector cartographic data which also possess some form of feature identification or tag. Recent research and development efforts at major commercial firms and (as well as DMA funded research, e.g. Rome Research Corporation - Automatic Cartographic Feature Identification II) are focused on producing such capabilities. Pattern/symbol recognition, template matching and artificial intelligence methodologies are being applied to enhance the current capabilities of raster-to-vector conversion software.

2.4.2.3 Algorithmic Approaches

Standard algorithmic approaches to the basic "dumb" conversion of raster structured cartographic data have evolved. The three basic steps are: skeletonization, line extraction and topology reconstruction.

2.4.2.3.1 <u>Skeletonization</u>.[9] Skeletonization subdivides into three approaches: peeling, ballooning and medial axis. Peeling refers to the reduction of each cartographic element by one unit of resolution at a time until only one unit of resolution remains. According to Donna Peuquet in her article "An Examination of Techniques for Reformatting Digital Cartographic Data: I - The Raster-to-Vector Process" (USGS), four aspects of this approach are worthy of note: 1) Execution time is a linear function of line thickness, 2) Peeling is sensitive to line weight variation resulting in irregular center-line definition 3) Thick or rounded lines cause difficulties in finding line centers and, 4) The algorithm is highly sensitive to overall line quality.

Ballooning is implemented where areas between lines are expanded until the lines separating them cannot be reduced further without causing a break in the boundary. According to Peuquet the following three points should be made about this approach: 1) This is the logical opposite of the peeling approach, 2) Node positions tend to be dislodged using this procedure 3) It gains speed by using the faster but less accurate four neighbor matrix search for bulk processing.

The medial axis approach essentially calculates centerline positions by assigning "nearness-to-edge" rankings to all pixel locations. Peuquet has four evaluations of this approach: 1) It is the fastest approach 2) It utilizes never more than four passes, 3) Its overall efficiency is affected by the number of pixels that must be looked at, thus line thickness and network density do affect the procedure and, 4) Errors can result from very thick nodes.

[9] Refer to Figure 3A for graphic representations of alternative skeletonization approaches

Figure 3A

Skeletonization

Peeling

```
X X X X          X X            X
X X X X          X X            X
X X X X          X X            X
 X X X X          X X            X
 X X X X          X X            X
  X X X X          X X            X
   X X X X          X X            X
```

Ballooning

```
0 0 X X X 0 0    X X 0 0 0 X X    X X X 0 X X X    0 0 0 X 0 0 0    X
0 0 X X X 0 0    X X 0 0 0 X X    X X X 0 X X X    0 0 0 X 0 0 0    X
0 0 X X X 0 0    X X 0 0 0 X X    X X X 0 X X X    0 0 0 X 0 0 0    X
0 0 X X X 0 0    X X 0 0 0 X X    X X X 0 X X X    0 0 0 X 0 0 0    X
0 0 X X X 0 0    X X 0 0 0 X X    X X X 0 X X X    0 0 0 X 0 0 0    X
```

Medial Axis (Adapted from Dr. Donna J. Peuquet, "An Examination of Techniques for Reformatting Digital Cartographic Data: I - The Raster-to-Vector Process", U.S. Geological Survey)

```
111      111         111      111        X       X
121      121         121      121        X       X
 1211   1232          1211   1221         X     X
  122  12343           121  12111          X   X
   121123               121121              X X
    12234                12211               X
   1233                 1221                 X
   12341                12221                X
  1223452              121121               X X
 123    121           121  121             X   X
 123    121          121   121            X     X
 123    121          121   121           X      X
 123    121          121   121           X      X
123     121         121    121           X      X
123     121         121    121           X      X
123     121         111    111           X      X
```

X = Bits turned on   0 = Bits turned off   1-5 = Pixel distance from edge (or closeness to 0 bits)

2.4.2.3.2 Line Extraction. Two basic techniques are available for line extraction: line-following and scan-line. In the line-following approach each individual line is followed from pixel to pixel, in any direction, until the end of the line is reached. According to Donna Peuquet, execution time is a linear function of the total length of lines to be vectorized. This approach also generates topological relationships by default. In the scan-line approach all lines intersecting a given scan-line are processed simultaneously and each scan-line is read only once. Peuquet asserts that a great deal of "bookkeeping" to track multiple lines simultaneously is required. More specifically, the overall line density and range of scan lines covered by individual map lines increase these bookkeeping overhead costs (i.e., time) as a linear function.

2.4.2.3.3 Topology Reconstruction.[10] The explicit definition of the spatial relationships of points, lines and areas in a cartographic data file usually occurs, to some degree, as a by-product of the skeletonization and line extraction procedures. However, the development of a fully defined topological data set usually requires more sophisticated post-processing of the vector data.

2.4.2.4 Speed

Speed is obviously an important factor in analyzing R/V routines. As mentioned above, different algorithmic approaches produce different speed results with different data types, densities, and conditions. Any comparison of R/V speed performance by various software routines must be made with cognizance of their different approaches and design functions. Basic "dumb" vs advanced "intelligent" processing is one critical factor influencing speed.

2.4.2.5 Quality

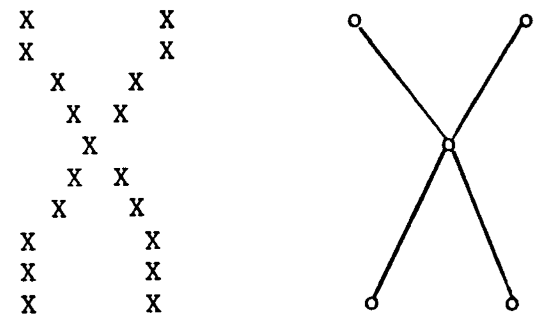
Vector data quality is equally as important as throughput speed. Large numbers of gaps, spikes and other anomalies in data which has been

[10]Refer to Figure 4A for graphic representation of Line Extraction and Topology Reconstruction

Figure 4A

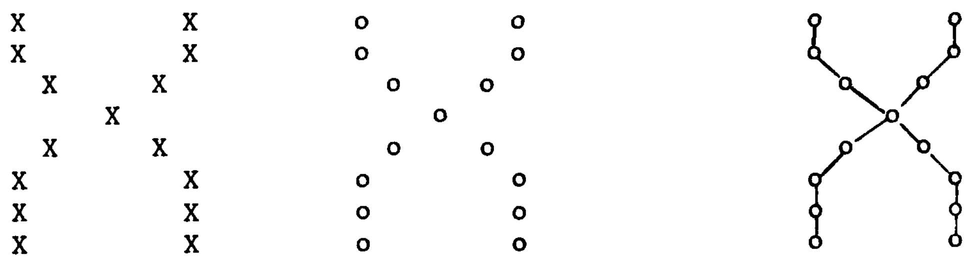
Line Extraction | Topology Reconstruction

Line Following



(Topological reconstruction is already accomplished via the line-following approach)

Scan Line



X = Bits turned on   o = x,y node coordinate positiion converted from a raster format to a center-line vector format result in large amounts of production resources being used to correct these problems. A balance must be found between speed of processing and maintenance of quality standards. This is partially influenced by the type of vector data editing capabilities which exist and their degree of sophistication.

2.4.2.6 Output Formats/Conversion Compatibility

The format of vector data which results from R/V conversion must be accessible to further data processing and manipulation. Organizations typically support in-house formats (e.g., DMA - Standard Lineal Format SLF). The ability to convert from a system vector format to the in-house format with relative ease and efficiency is critical. Many commercial vendors offer to write "custom-converters" upon request. Some provide standard conversions to various formats such as SIF (Standard Interchange Format - Intergraph) and plotting formats such as Gerber, CalComp, Applicon and Versatec.

2.4.3 Raw Vector Data Review/Edit

Vector data which results from raster-to-vector conversion typically exhibits various kinds of errors and anomalies. Facilities are available for the detection and correction of these problems. The functions performed and the techniques employed are roughly equivalent to those used in the raw raster data review/edit phase. (Please refer to section 2.4.1.2 Error Detection/Correction for further details). It should be noted that some error types are specifically attributable to the raster-to-vector conversion routines and some automated routines may differ from their raster counter-parts in those instances.

2.5 Analog-to-Vector Conversion: Tagging

Cartographic features which have been raster scanned (or digitized) require some form of feature identification or tag. This identification procedure is performed in different locations and times depending on which system of data capture is in use. Manual digitizing systems and automatic line-following systems, for example, provide for feature tagging as a concurrent process. The use of raster technology typically requires feature tagging as a post-processing function in vector mode.

2.5.1 Techniques and Procedures

Three approaches to feature identification and tagging are available or under development. These include: interactive, computer-assisted and automated. Interactive feature tagging usually is performed in vector mode through the facilities of computer graphics. Individual feature elements are identified through the use of light pens, cursors or trackballs and tagged with an appropriate code (alphanumeric). Actual code entry is accomplished through a keyboard or menu. Computer-assisted tagging results from human operator/computer interaction where, based on parameter inputs the "system" automatically assigns codes to a number of features. Some commercial firms offer routines for computer-assisted tagging of contours. Human operators identify starting contour values and "ending" contour values via a graphic transect, indicate a contour interval and the tagging is accomplished "automatically". A recent approach involves the use of benchmarks. Assigning of benchmark elevation values is sufficient to permit software tagging of all contours in a file. Spatial tagging software is also available whereby the boundaries of closed polygons are automatically identified as a result of operator identification of a position within the polygon. Automated feature tagging routines are currently being developed (as previously mentioned) for implementation in a raster mode or as part of the raster-to-vector conversion process. This latest development represents a powerful technological capability and if fully successful will revolutionize the analog-to-vector conversion process by greatly reducing the need for labor-intensive human interactive feature tagging.

2.6 Analog-to-Vector Conversion: Spatial Coding

The A/V Model Glossary defines Spatial Coding as consisting of three components: Universal Referencing, Topological Encoding and Data Structuring. Most commercial systems provide facilities to perform various aspects of these functions. In all cases it is unlikely that they will perform to user expectations or specifications. Typically, Universal Referencing is limited to less than a "handful" of map projection transformations. Topological Encoding will vary from system to system in terms of structure employed. Again, the best one can hope for is a topological structure which can be easily transformed or enhanced to meet in-house requirements. The same can be said for Data Structuring.

2.7 Analog-to-Vector Conversion Processing: Data Management

Data management is the logical and procedural conclusion of the analog-to-vector conversion process. In terms of raster-to-vector conversion evaluation little emphasis will be placed on this function.

3.0 ANALYSIS OF RASTER-TO-VECTOR CONVERSION*

The results of a study aimed at uncovering existing raster-to-vector conversion (RVC) algorithms and analyzing the suitability of these algorithms to DMA's overall analog-to-digital processing requirements are contained in this section.

The first step of the analog-to-digital conversion process involves the transformation of an optical cartographic image to electrical signals. Various equipment is available to perform this transformation. In considering RVC algorithms, it is most convenient to assume that the transformation is performed by a digitizing scanner. The digitizing scanner produces a finite sequence of digital information representing the image as a two-dimensional matrix of picture elements. This digital representation of the image is called a raster image because, similar to a television picture, it consists of numerous, thin, and parallel lines. In effect, each row of the matrix is directly mapped to a raster line. The digital raster image is a good starting point for investigating the RVC process because it represents a data format common to most RVC algorithms. The objective of RVC is to convert this raster representation of the map image to a sequence of vectors. It should be noted that the digitized raster image must be an accurate representation of the original image. In fact, it should even be better than the input. Map users have substantial tolerance to local degradations in graphic quality. It is contemplated that this tolerance is largely due to the fact that the human processes the data at a high level of abstraction. In contrast, machine processing of cartographic information is done at a very low level - often pixel-by-pixel. Because almost no contextual abstraction is possible at this low level of processing, machine processing of cartographic data is extremely sensitive to graphic quality variations.

* All footnotes in this chapter refer to the bibliography at the end of Section 3.0.

As an example, consider a road symbol which has a circular defect approximately in the center (see Figure 5A). A reasonable vectorization of this road, obtained through low - level processing, is also shown in Figure 5A. Note that even with the small amount of contextual information, a human could effortlessly detect the error and create a correct vector representation for this road.

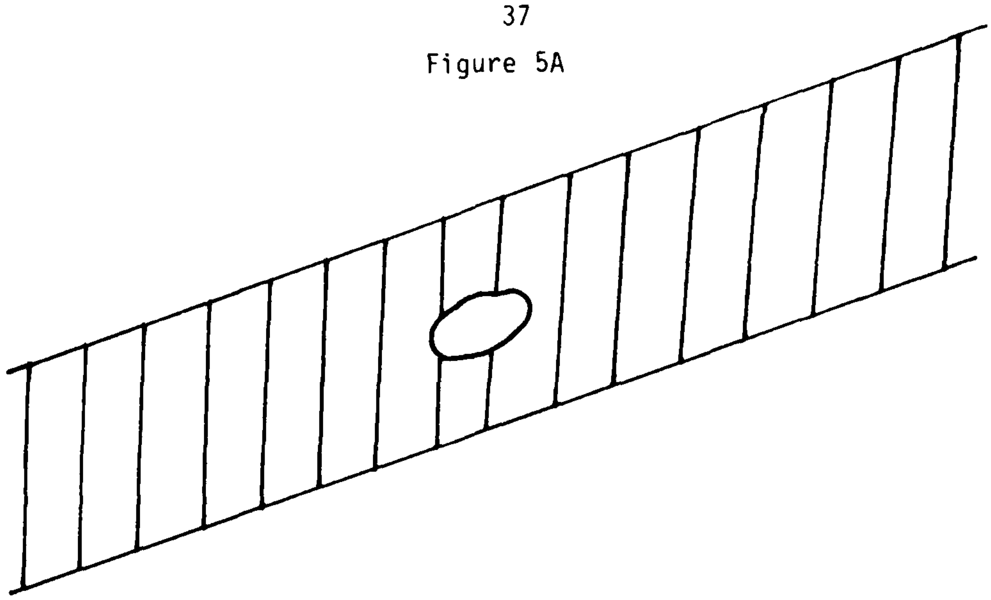
To overcome problems with graphic quality of the input document and other sources of noise, additional processing is often performed. This processing takes the form of filtering. A number of techniques have been described[1] and a common one is a moving-window average. It should be noted that these filtering techniques are only partially effective, primarily because they use statistical information contained in a large number of adjacent pixels. However, they fail to fix all problems because high-level abstractions of the map are not formulated.

RVC is an important component of the overall analog-to-digital conversion process. Raster-to-vector conversion remains essential because a large number of algorithms for automated processing of cartographic data have been developed for maps represented in vector formats, and counterpart algorithms for raster encoded maps have not been developed, The optimal routing problem[2] is an example of a problem for which only a vector oriented algorithm exists.

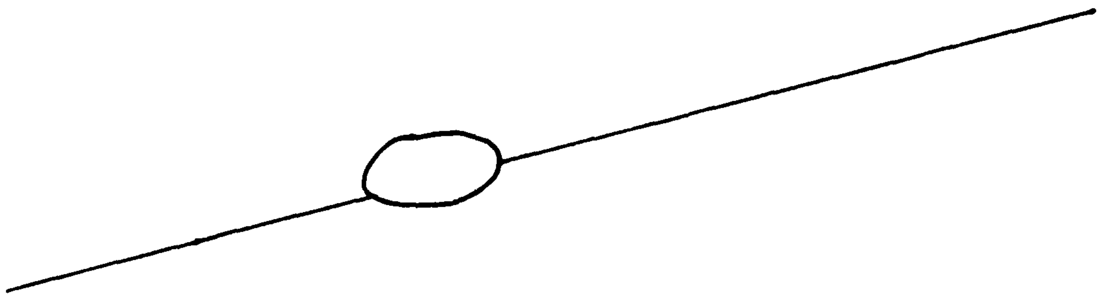
3.1 Some Algorithms for Raster-to-Vector Conversion

RVC is the process of transforming raster image data to vector data. Although numerous equipment and software for performing RVC exist, the number of unique conversion algorithms is not as great. Most of the algorithms described in the literature are variations on a few basic themes, and the raster-to-vector conversion software and equipment investigated display a great similarity in the algorithms used.

Figure 5A



Road symbol with a graphic error.



A reasonable vector representation of the above road obained through low-level processing.

The first distinction encountered between RVC algorithms is the data format which the algorithms process. It was stated earlier that a starting point for RVC is the raster image. Some algorithms process this pixel data repeatedly to produce the appropriate vector image representation. Other algorithms require the raster image data to be converted to an intermediate representation, such as an XY boundary sequence, prior to vectorization. An XY boundary sequence is the image contour defined by a sequence of eight-connected or four-connected pixels[1].

Most of the published works describe RVC algorithms which process pixel data. The development of boundary-chain processing algorithms is more recent, and publications dealing with this RVC approach are not as common. This report treats the two RVC techniques separately in the following sections.

3.1.1 Pixel Processing RVC

Within the set of algorithms which process pixel data, another distinction is encountered. Some of these algorithms follow each map line until the entire map sheet has been vectorized (i.e., line-following), while others perform the vectorization by processing raster lines (scan-line processing). To better understand how line-following differs from scan-line processing, it is convenient to decompose the RVC process into the following operations:

- o Map Segmentation - the map is segmented such that each map segment can reside in physical memory.
- o Line Thinning - map lines are reduced in thickness to a single pixel width.
- o Vectorization - the extraction of lines contained in the thinned raster image.
- o Topology Reconstruction - joining of map segments and identification of vector intersections and nodes.

The above operations can be implemented sequentially (in the above order) or several of the above operations can be performed simultaneously. It is often desirable to perform simultaneous processing since a time saving may result.

For the benefit of clarity, the RVC process will be described in terms of the above operations as if they were executed sequentially. Finally, techniques used to improve the performance of conversion by circumventing sequential processing will be discussed.

3.1.1.1 Segmentation

The objective of the segmentation operation is to improve efficiency of the conversion process by reducing the number of accesses to slow storage devices (e.g., disk or tape). The objective is accomplished by dividing the image data into pieces which fit into the computer's physical memory. Each piece is then processed separately. Ultimately, the results of each processed segment are joined into a complete map image.

Segmentation provides substantial benefit for algorithms which make more than one pass through the raster data. Algorithms which make a single pass may also benefit from segmentation. Single-pass scan-line processing algorithms[4,5] have been described. These algorithms output vector data as each raster line is processed. If the quantity of vector data exceeds the amount which can be stored in physical memory, then secondary storage has to be used resulting in longer data access times.

Algorithmically, the segmentation operation is quite simple. Pieces of the raster image are copied into logical data entities (typically files) of the appropriate size. The map segments can be either patches or scan-line swaths, whichever is more convenient. The segmentation software may also perform a data format transformation. The input raster data can be encoded in a variety of ways. Examples include (1) raw raster data consisting of a bit-stream where bits correspond to on/off pixels or a byte-stream where bytes correspond to pixel grey-levels or colors and (2) compressed raster data such as run-length encoding. During the segmentation operation, the input data can be easily transformed to the format required by subsequent operations.

In and of itself, the segmentation operation will not produce any errors. However, its interaction with the previous processing may indeed cause problems. Specifically, any lines or features on the segment edge may be subject to error. Consider for example a line which is coincident with a segment edge and parallel to that edge. If the line is more than one pixel wide, then portions of it will be present in each segment. Subsequent thinning and vectorization will produce a vector corresponding to this line in each segment. Ultimately, two vectors slightly offset result from a single line. To overcome such problems, it is possible to overlay the segments by at least twice the width of the thickest line, and ultimately, join the segments at the center of the overlap region.

3.1.1.2 Line Thinning

The line thinning operation is the most time-consuming and error-prone operation in the RVC process. Line thinning algorithms can make numerous passes through the image data resulting in slow operations. Additionally, thinned output is often considered erroneous because unanswered questions remain concerning the correct thinned representation of map features.

Line thinning is often referred to as skeletonization. The typical definition of a skeleton[1] identifies the locus of points which are centers of maximum diameter circles contained within the raster image. Given this definition, the following two problems need to be addressed (1) How is such a skeleton obtained? and (2) Is this skeleton required for cartographic processing?

Pixel-processing skeletonization algorithms have been summarized, analyzed, and compared extensively in the literature[6,7,8]. Most of the skeletonization algorithms move a small submatrix of pixels (called the decision matrix) within the whole raster image. As the submatrix is moved through the image matrix, set pixels are cleared (i.e., black map areas are changed to white) if certain rules are met. The rules are primarily selected to fulfill the following criteria: (1) maintain connectivity of lines, (2) retain endpoints of lines, and (3) eliminate directional bias of thinning. The rules may be expressed in a variety of ways, such as table look-up operations[9] or bit-wise (i.e., decisions), while the third criterion requires that the image be processed in alternate directions (or perhaps that the decision matrix be rotated[8].

With regard to the second question, present experience indicates that the skeleton described by the above definition is quite useful for extracting the characterizing elements of a given pattern and reconstructing the original pattern from its skeleton, however, its applicability to automated cartography is questionable. The difficulty with using the skeleton for cartographic applications is that numerous branches tend to be generated which are contained entirely within the original map line. Although quite useful for describing the shape of the original image pattern, in cartographic processing these branches ultimately result in stubs (see Section 2.4.1.2.1 for a discussion of error types). This difficulty is typically overcome by additional processing such as filtering and editing. Another difficulty with skeletonization for cartographic processing is that the true midpoint of a cartographic feature tends to be biased at feature intersections (e.g., road intersections). Again, this problem can be overcome with additional processing. A specific algorithm to overcome this problem is presented in the section dealing with non-pixel processing RVC algorithms (Section 3.1.2).

At present, no thinning algorithms appear to be ideally suited to cartographic processing. In order for a thinning algorithm to perform desirably in the context of RVC it must fulfill cartographic requirements and not necessarily image processing or feature extraction requirements. It must be remembered that production RVC systems use people for feature tagging. The objective of the line thinning is to reduce all lines to a single pixel width while retaining the location of the line center. The shape of the lines must also remain unchanged. Algorithms which have been developed to generate skeletons applicable to cartographic features may be useful if it is desired to perform automatic feature tagging during the analog-to-digital conversion process. To date, automatic feature tagging has been attempted by processing raster data[2,3] with some success. However, an attempt to recognize thinned and vectorized feature data may prove to be easier. An attempt leading in the direction has already been reported[11].

3.1.1.3 Vectorization

The extraction of lines contained in the thinned raster image is typically provided as a chain-encoded line structure[1]. During this task, the distinction between line-following and scan-line processing becomes most obvious.

3.1.1.4 Topology Reconstruction

The purpose of topology reconstruction is to undo what has been done by segmentation and to provide information of the feature network represented by the map.

3.1.2 Non-Pixel Processing Raster-to-Vector Algorithms

The pixel processing algorithms previously described tend to modify the XY image phase based on repeat inspection of small sub-matrices of the image. Non-pixel processing RVC algorithms tend to deal with the image content at a higher level of abstraction (e.g., map-line boundaries). This higher level view of the map features requires fewer data accesses to compute the centerline of a thick map feature. Furthermore the computations required are a relatively straightforward application of plane geometry concepts.

Several vectorization systems using non-pixel processing algorithms have been reported[12,13,14]. In one of the recent works[4], the XY locations of feature edges are used to determine the location of the features center. The concept of the line adjacency graph (LAG) has also been presented[1,2] and applied to vectorization. The concept of symmetric axis transform (SAT)[14] may also be applicable, but probably requires modification for specific needs of RVC, which is not necessarily feature extraction. Of particular interest is the work performed on the Minimum Base and Equal-Diagonal algorithms[13]. Aside from being computationally efficient, and impervious to stub generation, the algorithm presents a possible solution to the T-intersection error. The authors interpolate the edges of the feature essentially cutting off intersecting features. Centerline calculations are performed for each feature separately and later joined.

Another non-pixel processing algorithm involves the analytical propagation of wavefronts from a features edge[15]. However, this approach does not appear to have computational advantage, and it still suffers from stub and T-intersection errors.

3.2 Raster-to-Vector Conversion Implementation

Analog-to-vector conversion may be accomplished in a variety of ways. Equipment and software for manual vectorization are quite prevalent. As was mentioned earlier, various automated systems for converting digital raster images to vector data are also commonly available, This section of the report focuses on the organization of the electronic circuitry used to implement the RVC algorithms discussed in the previous sections. This discussion is limited to automated RVC systems apart from operator-assisted ones because it is believed DMA production requirements will primarily be fulfilled through the use of automated RVC systems.

3.2.1 Background

The initial form of the cartographic product is a visible image. Unfortunately, present technology does not allow processing of a map in this form. The image must be transformed into electronic information for processing. This transformation is accomplished by digitizing the optical image. After digitization, the image is represented as an array of 1s and 0s. (for simplicity, a two level black-and-white image is assumed.)

Each binary image pixel can be easily stored and manipulated in electronic circuitry. This circuitry is described in numerous introductory texts on digital electronics[16,17]. Basically, the storage elements are called memory cells or flip-flops, and each cell is able to store one bit (pixel). Pixels are manipulated with logic gates which perform Boolean operations such as AND, NOT, and OR. By combining these fundamental Boolean operations, more complex logical operations can be realized such as implication (IF-THEN), equivalence, and exclusive-OR. The output of these logical operations can represent a higher level of abstraction of the input raster image data. These higher level abstractions can be stored in memory cells as finite-length binary codes.

The memory cells and logic gates provide the building blocks for the vectorization system. Given the ability to store and manipulate image pixels a method of controlling the processing is still required. That is, the ability to selectively manipulate certain groups of pixels and allow results of an operation determine what additional processing is required. Therefore, the concept of instructions is necessary. We now have the three elements required to implement a vectorization (or any data processing) algorithm. To summarize, these three elements are data storage, operational hardware, and control hardware. Note that control hardware is also composed of memory cells and logic gates. A fourth element, input/output is also needed for obtaining and displaying the data. Even though it does not directly impact the algorithm's implementation I/O bound algorithms can be very ineffectual. For example, the Goodyear Staran R/V processor is reported to have been I/O bound thus dramatically reducing its effectiveness.

3.2.2 Various Processing Architectures

It is possible to organize the storage, operational control and I/O elements previously described in a variety of ways and accomplish the same processing results. The following subsections describe several techniques which are common.

It was earlier mentioned that instructions were used to specify the processing steps. In practice, these instructions may take several forms: (1) program statements, (2) microinstructions, and (3) logic gates. The distinction between the instruction forms provides a convenient organization for discussing alternative algorithm implementations. These alternatives are (1) software (2) micro-programming, and (3) hardware.

It should be noted that the method of implementing a specific algorithm (e.g., software or hardware) will in no way influence the quality of the output. Accuracy may be influenced however, as determined by wordsize for example. Time required to obtain the same results may vary orders of magnitude between implementations. Thus speed, not quality, is the subject of this discussion.

3.2.2.1 Software Implementation

The great majority of vectorization algorithms have been implemented in software. The software takes the form of programs which consist of very specific, low-level instructions. The programs represent a translation of the algorithm into a collection of instructions which a computer can perform. The computer itself is actually an implementation of the following simple algorithm:

```
Do forever:
Fetch the next instruction;
Execute the fetched instruction;
Check for special events;
End do.
```

Numerous texts are available which discuss the actual implementation of general purpose computing machines[18,19]. This discussion will focus on the limitations imposed by the computer on the implementation of vectorization algorithms.

Software implementation is by far the slowest possible implementation of those considered. Several factors cause this: (1) Instruction set is general and not specifically suited to vectorization resulting in a large number of sequential instructions being required, (2) Data path between CPU and storage is of limited size requiring numerous data transfers, and (3) Memory organization is not suited to storage of a large two dimensional matrix and more importantly, to the retrieval of a small submatrix of this matrix. The third problem is especially time consuming for pixel-processing algorithms.

Within the scope of software implementation, techniques exist for substantial improvements in performance. These improvements essentially result in the system being able to perform more (or more complex) operations. A technique which has been reported in the literature is array processing[20]. The results of the Goodyear STARAN research, for example did not meet expectations primarily, we believe, due to the systems I/O constraints.

Recent advances in computer organization suggest that multiprocessing systems (especially multiple instruction multiple data organization) will substantially improve performance of software implemented image processing algorithms. Some work has already been reported in this area[21], but additional research is required before this technology finds widespread use.

Although the software implementation results in the slowest performance, it is the most popular for an obvious reason; algorithms are easy to implement and change. The algorithms are implemented with alphanumeric text, and as a result they are much more flexible than hardware. Software vectorization will continue to be widespread because of its relative ease of implementation and modification. In fact, software is the only reasonable choice for identifying the quality related characteristics of a new algorithm. However, once the quality of processing has been proved, implementation other than software can provide substantially improved performance.

3.2.2.2 Microprogrammed Implementation

Microprogrammed control is a well established digital technology, and much literature exists decribing the fundamental concepts[22,23]. In fact, many modern computers are implemented using microprogrammed technology.

The advantage of this implementation is the ability to include instructions and data organization specific to the algorithm being implemented. The result is that a few micro-instructions could perform the same processing that tens of software instructions might do.

Although unreported, some work has already been done in this area with regard to the vectorization algorithm. Several microprogrammed computer implementations (e.g., DEC VAX-11/780 and HP1000) permit user access to the microprogram. It is possible for the user to include his own special instructions. At least one manufacturer of vectorizing equipment has taken advantage of this capability.

3.2.2.3 Hardware Implementation

Overall, microprogrammed implementation can provide substantially improved performance over software implementation. However, a hardware implementation utilizing vendor logic interconnection can provide the best performance. Several methods of interconnecting hardware logic to obtain an algorithmic machine are possible. These include synchronous and asynchronous operations organized sequentially or in parallel. In fact, various combinations of the above are possible. Methodologies for the transformation of algorithmic descriptions to hardware have been developed[24]. One of the degrees of freedom in the hardware is the allowed level of parallelism which greatly influences the performance of the implementation.

Ongoing development in algorithmic state machines, VLSI technology, and silicon compilers will provide a vehicle for rapid implementation of hardware vectorization systems.

Bibliography

1. Pavlidis, T., Algorithms for Graphics and Image Processing, Computer Science Press, Rockville, Md. (1982), pp. 195.

2. Gillotte, M. J., "Raster Processing Techniques", Final Technical Report from Rome Research Corporation to Rome Air Development Center (February, 1979).

3. Douglas, A. V., Edwards, T. V., Gillotte, M.J., and Rapke, M. L., "Litho Scanning Techniques", Final Technical Report from Rome Research Corporation to Rome Air Development Center (September, 1981).

4. Crane, Robert K., "Mini Raster-to-Vector Conversion", Final Report from Environmental Research and Technology, Inc. to US Army Engineer Topographic Laboratories (September 1981) ETL-0269.

5. Chakravarty, I., "A Single-Pass, Chain Generating Algorithm for Region Boundaries", Computer Graphics and Image Processing 15, pp. 182-193, Academic Press Inc., 1981.

6. Peuquet, D.J. And Boyle, A. R., Raster Scanning, Processing, and Plotting of Cartographic Documents, SPAD System Ltd., 1984 Williamsville, NY.

7. Davies, E. and Plummer, A., "Thinning Algorithms: A Critique and a New Methodology", Pattern Recognition 14 (pp. 49-72) 1979.

8. Lamura, II., "A Comparison of Line Thinning Algorithms from Digital Geometry Viewpoint", Proceedings of the Fourth International Joint Conference On Pattern Recognition, (pp. 715-719), 1978.

9. Mueller, C. S. M., and Christensen, A. H. J., "SCANLINE: A Computer Program for Converting Raster Maps into Line Maps", National Research Institute for Mathematical Sciences, Pretoria, South Africa, CSIR-TWISK-43, 1978.

10. Radosevic, R. G. et al, "Associative Array Processinq for Topoqraphic Data Reduction", Final Technical Report from Goodyear Aerospace Corporation to US Army Engineer Topographic Laboratories, ETL-CR-74-20, 1974.

11. Loew, M M., Pickholtz, R.L., and Goldman, L. "Image Processing for Cartographic Applications", 1981 Machine Processing of Remotely Sensed Data Symposium. (pp. 447-453).

Bibliography
(Continued)

12. Pavlidis, T., "A Minimum Storage Boundary Tracing Algorithm", IEEE Transactions on Systems, Man, and Cybernetics, V. SMC-8, N. 1, January 1978, (pp. 66-69)

13. Shapiro, B., Pisa, J., and Sklawsky, J., "Skeleton Generation from X,Y Boundary Sequences", Computer Graphics and Image Processing, V. 15, 1981 (pp. 136-153).

14. Blum, H. and Nagel, R., "Shape Description Using Weighted Symmetric Axis Features", Proceedings, IEEE Pattern Recognition Conference.

15. Montanari, U. "Continuous Skeletons from Digitized Images", Journal of the Association for Computing Machinery, V. 16, N. 4, October 1969 (pp. 534-549).

16. Williams, Gerald E. Digital Technology 1977 Science Research Associate, Chicago.

17. Mauo, M.M., Computer System Architecture 1976 Prentice-Hall, Inc. Englewood Cliffs, N.J.

18. Hamacher, V. C., Vranesic, Z. G., and Zaky, S.G. Computer Organization 1978 McGraw-Hill Book Co., NY.

19. Baer, Jean-Loups, Computer Systems Architecture, 1980 Computer Science Press, Rockville, Md.

20. Goodyear Aerospace Corp., "Associative Array Processing of Raster Scanned Data for Automated Cartography II," Report for U.S. Army ETL, 1977.

21. McDonald, W.C., and Smith, R.W., "A Flexible Distributed Testbed for Real-time Application", Computer October 1982, V.15, N.10, IEEE.

22. Myer, G. J., Digital System Design with LSI Bit-Slice Logic, 1980 John Wiley & Sons, NY, NY.

23. Kraft, G. D., and Toy, W. N., Microprogrammed Control And Design of Small Computers, 1981 Prentice-Hall, Inc., Englewood Cliffs, N.J.

24. Thayse, A. "Synthesis and Asynchronous Implementation of Algorithms Using a Generalized P-Function Concept", IEEE Transactions on Computer, October 1984, V. C-33, N.10, (pp. 861-868).

4.0 DMA ANALOG-TO-VECTOR CONVERSION PROCESSING (CURRENT)

Chapter 4.0, pages 51 through 73 inclusive, is published in a separate appendix entitled "Special Appendix to the Defense Mapping Agency (DMA) Raster-to-Vector Analysis". Distribution is limited to the Department of Defense.

5.0 DMA ANALOG-TO-VECTOR CONVERSION PROCESSING (FUTURE REQUIREMENTS)

Chapter 5.0, pages 74 through 78 inclusive, is published in a separate appendix entitled "Special Appendix to the Defense Mapping Agency (DMA) Raster-to-Vector Analysis". Distribution is limited to the Department of Defense.

6.0 AN EVALUATION OF STATE-OF-THE-ART ANALOG-TO-VECTOR CONVERSION SYSTEMS

Battelle has completed a topical evaluation of state-of-the-art A/V conversion systems. It is topical in the sense that the descriptions and evaluations are based solely on published reports, commercial literature, academic research, personal experience and conversations with knowledgeable people. No rigorous testing or benchmarking has been performed. A comprehensive literature search has been completed and a substantial amount of material has been collected. Compare and contrast charts for the systems under consideration have been compiled based on the A/V model developed in the earlier stages of this investigation.

6.1 Information Search

Battelle has compiled four separate lists representing the results of information searches (see Appendix C). The lists include: a) a bibliography; published papers about data acquisition systems, raster-to-vector conversion, analog-to-vector conversion, and related topics, b) a professional contact list; a list of all professionals active in related fields and endeavors who have been contacted, c) a commercial contact list; a listing of all commercial firms, and their representatives providing services and systems in the analog-to-vector conversion are [illegible] have been contacted (with a bias towards raster scanning/automatic line- [illegible] systems).

6.2 Compare/Contrast Charts

These charts have been developed to permit a concise cataloguing of information about state-of-the-art A/V systems to facilitate ease of comparison. The charts are designed according to the major components of the A/V conversion model. Information provided on the charts (derived from System Profile Questionnaires - see Appendix D) represents data which Battelle feels

[11] Compare/Contrast Charts are appended to the end of Section 6.0 (Figures 7A through 7E).

is valid and unambiguous. Data which remains questionable is indicated as such on the final charts.

6.3 DMA Applicability

Battelle has reached some conclusions about the potential applicability of various analog-to-vector conversion systems to the Defense Mapping Agency's digital cartographic production requirements. This section includes a brief summarization of each system under consideration, a section on highlights, a section on current research and development, and a DMA applicability statement. The following systems are reviewed: Scitex Response 280, MBB/Kongsberg SysScan, Intergraph/Optronics Scan Data Capture, Laserscan Lasertrak, Broomall Scan Graphics Automated Graphic Digitizing System (AGDS), ANA Tech Vana System, Teledyne Geotronics Linetrac, and Gerber Scientific VDS-2500 Digitizing System.

6.3.1 Scitex Response 280

System Summarization: a primarily raster-based turnkey system which provides facilities for all components of the analog-to-vector conversion process. Preparation - source input requirements are quite flexible with large format specifications and up to twelve colors permissible for each manuscript. Most common cartographic source materials are acceptable. Complex manuscript preparation and reformat/enhancement may be advisable to take full advantage of the color raster scanner's capabilities. Digitization - two types of scanners are available: a 40" X 36" electro-optical raster scanner with twelve color recognition and up to 1200 lines of resolution per inch, and a 42" X 75" laser raster scanner with sixty-four levels of grey level recognition and up to 2000 lines of resolution per inch. Raster-to-Vector Conversion - includes a package of interactive, computer-assisted and automatic raster data editing and manipulation facilites. Raster-to-vector conversion software has been reported[12] to demonstrate improved throughput and higher quality standards

[12]According to USGS Research Cartographers.

over the past few years. Raw vector data editing facilities parallel its raster editing counterparts. Tagging - in recent years Scitex has laid claim to development of interactive, computer-assisted and automatic feature tagging procedures. Of particular interest has been the computer-assisted/automatic elevating of contours. Demonstrations of this technique have been made at trade shows but further benchmark testing is recommended. Spatial Coding - three transformation routines are currently available for implementation in either raster or vector mode. According to Steve King, vice president of CAD/CAM Systems for Scitex America, an additional eight projection transformations have been promised to DMAAC pending award of a system acquisition contract. Support of the DMA Standard Linear Format (DMASLF) has also been commited to DMAAC in conjunction with a pending award. Data Management - internal file management procedures are provided, however, no spatial data management facilities are available.

System Highlights: These include a large format raster color scanning system, automatic raster data editing facilities, computer mapping routines, a large format laser film plotter, and good overall system integration. System deficiencies include: limited addressable CRT image, small format manual vector digitizing pad, mediocre system/function documentation (lacking cartographic orientation), and lack of direct electronic raster editing CRT hardcopy output.

Research and Development: A major research and development thrust involves advanced procedures for raster data processing. The development of automatic symbol/pattern recognition algorithms is being pursued. Extension of current automatic feature tagging capabilites (e.g., contour elevating) is being explored. The recent availability of the Scitex ELP large format raster scanner/plotter represents a significant hardware advance. A larger raster color editing CRT and a more powerful CPU are under consideration at this time (although no immediate movement in these two areas is currently projected).

DMA applicability Statement: The Scitex Response 280 system has a significant role to play in supporting the Defense Mapping Agency's digital cartographic production requirements. Its predecessor, the Response 250, is currently providing DMA support in the areas of color separation scanning, hydrographic chart compilation and revision, computer mapping, and topographic slope mapping. The expansion of system support for other topographic/hydrographic mapping applications is under consideration. Thus, the continued use of Scitex systems and technology is expected to continue for the foreseeable future.

6.3.2 MBB/Kongsberg SysScan System

System Summarization: a comprehensive digital cartographic data/map production system providing facilities in all aspects of the A/V conversion process. Preparation - flexible analog material input requirements, an effective limitation on number of colors per manuscript (2-3) and a reduced format specification. Digitization - state-of-the-art solid state raster scanning (23.6x39.3" flatbed) with variable resolution up to 1000 lines per inch and limited color recognition (strong grey level recognition, however). Raster-to-Vector Conversion - parameter controlled realtime or batch raw raster data editing software routines are available. Raster-to-vector conversion contains a special test pattern facility and hardcopy electrostatic output device for visual review. Raster-to-vector conversion is accomplished via centerline and/or edge line (boundary) determination. Both are parameter controlled for filtering and accuracy. A full interactive vector data editing hardware/software system is also provided. Tagging - complete interactive tagging facilities with a computer-assisted/automatic contour tagging routines are available. Spatial Coding - spatial coding routines are provided to structure polygon boundaries by center-of-polygon identification methods. Some universal referencing routines are available. SysScan supports vector polygon data structures and is especially adept at producing digital terrain model data (according to DMA specifications). Data Management - a hierarchical attribute data management system is provided.

System Highlights: These include efficient raster scanning, automatic raster filtering/editing routines, full vector data display, edit and manipulation facilities, computer mapping techniques and good system integration. Sysscan has also recently introduced a full raster editing capability, DatEdit. System deficiencies include limited analog manuscript format size and single pass color scanning recognition.

Research and Development: According to Dr. Gunter Oesterhelt, MBB-SysScan Technical Director, key development areas include: combined raster/vector hardware-software (i.e., DatEdit), larger format color raster scanners (SysScan will soon announce two new scanners, FastScan 19 up to 19" wide and FastScan 47 up to 47" wide with endless length), and sophisticated automatic feature identification software. (Recent developments claim a symbol/pattern recognition library for handwritten numbers of DTED/DFAD like polygon-feature code data.)

DMA Applicability Statement: This complete system appears to provide capabilities which are compatible with DMA's production requirements. The system would be very effective in DTED generation and related elevation data production. Its increasingly strong emphasis on raster processing with an integrated, fully complemented vector system represents state-of-the-art capabilities. Further investigation is warranted.

6.3.3 Intergraph/Optronics Scan Data Capture System

System Summarization: a comprehensive raster scanning, vector processing turnkey cartographic data capture system recently developed by Intergraph Corporation. Preparation - flexible analog source documents requirements including large format size and most typical materials. Digitization - two types of scanners are available: PLT 125 Scanner/Plotter System and PLT 126 - Scanner System (40x40"). Scanning resolution up to 1000 lines per inch is supported along with black and white, continuous tone and color document scanning. Color scanning requires three scan passes through red, green and blue filters. Raster-to-Vector Conversion - Limited raw raster data review/edit processing. Raster-to-vector conversion is performed on an Intergraph Graphics Processor and pipelined to a high resolution color interactive work station for further processing. Automatic, computer-assisted, and interactive procedures are available for vector review/edit. Optimization of automatic processing with "queued edit" scenario support for human operators is a system priority. Tagging - Intergraph provides automatic functions for symbol/character recognition and feature tagging. This is performed in conjunction with "queued edit" scenarios for human operator decision-making optimization. (It is unclear at this time as to the level of automatic recognition/tagging success which is currently system supported.) Spatial Coding - interfaces with link-node and chain-enclosed polygon data structures along with Bureau of the Census "Dime", U.S. Geological Survey "DLG" and Defense Mapping Agency "SLF" structures. A package of twenty map coordinate transformation routines is also provided via the World Mapping System (WMS).

System Highlights: These include a large format/high volume raster scanner, procedures for automatic character/symbol recognition, and queued edit scenarios to optimize human operator interaction. Deficiencies include the need for multiple pass color raster scanning.

Research and Development: According to David Sinton, senior Intergraph Technical Manager, Intergraph is dedicated to developing procedures for the capture of analog cartographic documents which fully automate the conversion of line drawings, symbols, characters and photographic images. The emphasis on developing procedures and methodologies for optimizing human operator interaction also continues.

DMA Applicability Statement: Based on documentation and personal conversations with Intergraph personnel as well as experience with other Intergraph systems, it appears that the Scan Data Capture System has a potentially significant role to play in fulfilling DMA cartographic data production requirements. Their emphasis on advanced "intelligent" data processing could alleviate significant labor intensive procedures currently implemented at DMA.

6.3.4 Laserscan Lasertrak

System Summarization: Laserscan Lasertrak offers a total digital cartographic production system, representing the only automatic line-following system to be fully evaluated. Preparation - severe material preparation requirements are imposed on the user. Lasertrak only digitizes a small format (3.84" x 2.7" image) negative image necessitating reprographic work prior to digitization. Digitization - deflected laser local automatic line-following. Tagging - facilities are provided for feature tagging concurrent with the digitization process. Spatial Coding - an internal feature format (IFF) results from digitization. Routines are provided for universal referencing to a number of international map projections. Custom software can be developed to provide data structure conversion routines to match user requirements. Editing - a package of interactive editing routines is available to correct any errors or anomalies resulting from digitization or poor input materials. Data Management - spatial attribute data base management facilities are provided to facilitate editing or map production.

System Highlights: Laserscan Lasertrak provides facilities for "intelligent" digitization, with resulting data residing in a flexible, accessible vector, feature-tagged data format. The system is particularly efficient in the capture of contour data (and similarly, generic non-intersecting lines) and closed polygon data. Difficulties have been reported in effective data capture of dense intersecting networks. Obvious difficulties result from cartographically symbolized data input. The requirement for photo-reduction and negative image production of input documents is another drawback.

Research and Development: Laserscan Laboratories, located in Cambridge, England, maintains a staff of researchers and facilities for system testing and development. No information is currently available about recent research and development trends at Laserscan. Two systems have recently been delivered to the U.S. Geological Survey. A visit to their production facility might provide first-hand information on the state-of-the-art. (These systems are currently not in production).

DMA Applicability Statement: The Laserscan Lasertrak system may provide effective service to the Defense Mapping Agency for the capture of contour data (DTED and related products) and possibly for polygon products such as DFAD and polygonal DTAD. The advantages of this approach (compared to raster technology/raster-to-vector conversion) must be weighed against material input requirement limitations and increased human operator dependence.

6.3.5 Broomall Automated Graphic Digitizing System (AGDS)

System Summarization: a turnkey raster scanning, interactive vector processing cartographic data capture system providing basic procedures for all components for the analog-to-vector conversion process. Preparation - large format analog source material acceptability limited to black and white or single color (no red) representation. Digitization - large format "red-light" laser flat-bed raster scanning with single color recognition and scanning resolution of 1000 lines per inch. Raster-to-Vector-Conversion - raw raster data review (only), software for raster-to-vector conversion, and interactive computer graphics for raw vector data review/edit. Tagging - interactive computer graphics for human operator feature tagging. Spatial Coding - facilities for table coordinate to geographic coordinate transformation. Data Management - internal file management procedures are provided.

System Highlights: A basic work-horse data capture system representing a known quantity to DMAHTC and DMAAC. Its deficiencies include: black and white scanning (only) at a fixed resolution, no raster editing. Recently, Broomall has developed an improved scanner camera providing for .002" spot size fidelity. In addition, its realtime rasterization/vectorization (RAVE) system also represents an attempt to improve data capture throughput. This is somewhat diminished by the lack of any raster data editing facilities.

Research and Development: No reported research and development.

DMA Applicability Statement: This system continues to meet basic DMA data production requirements. Its inability to address errors and anomalies associated with this process in more automatic ways is a serious handicap. It appears that for DMA to meet the increasing demands of the future, the AGDS will require further upgrades or replacement.

6.3.6 ANA Tech Vana System

System Summarization: This is an example of a proprietary hardware raster-to-vector conversion system which integrates with a number of commercially available raster scanners and computer systems. Preparation - analog source materials are limited to black and white (or single color) representation whose format acceptability depends on individual scanner limitations. Digitization - this vectorization system currently links to the following commercial scanners: Crossfield, EOCOM, ECRM, Imagitex and Industrial Vision Systems. These vary in size from 12" X 17" to 36" X continuous length. Their scanning resolutions vary from 200 to 1400 lines per inch. Raster-to-Vector Conversion - a proprietary process which runs concurrent with raster scanning. According to conversations with ANA Tech engineers, R/V conversion is accomplished in two steps. First, a hardware process converts, in real time during the raster scanning process, all relevant area boundary chains, retaining all connectivity information. Parameters include: thresholding (how closely to model lines), vector smoothing, fuzz elimination, 2D hole filling and line prediction across broken lines. Straight lines are not segmented due to any scanning banding limitations.

This is followed by a software process which determines the vector centerline information. This process populates a dynamic graphical data base which retains all the vector centerline chains with line weight information, all area information, all connectivity information, has layering capability, and has the capability to add intelligence to all graphical information, e.g., character/symbol recognition. During the population of the data base additional fuzz elimination, smoothing, and snow elimination can be performed.

Additional interactive software functions include: Character and symbol recognition, graphical editor (used to cleanup and add additional intelligence) and output formatting (the capability to output all relevant information from the data base to the end user system).

Future applications will consist of text recognition, curvature analysis, recognition of mathematical entities, and automatic layering based on line weights.

Lineweight tagging is also performed during R/V conversion. Tagging - no feature tagging software is available at this time. Spatial Coding - no spatial coding facilities are available at this time. System Highlights: Very fast (according to company PR) proprietary hardware for raster-to-vector conversion. Symbol and character recognition is also claimed. Deficiencies include an inability to process multiple color input and a basic lack of cartographic orientation.

DMA Applicability: The ANA tech system does represent an alternative option to DMA to configure a unique cartographic data capture system as opposed to acquired "packaged" systems. It is still unclear whether this hardware R/V conversion process is effective for cartographic applications. However, Edward Getchell, president of ANA tech, has indicated a strong interest in working with potential cartography clients in developing application methodologies more relevant to their unique requirements.

6.3.7 Teledyne Geotronics Linetrac

System Summarization: Proprietary hardware raster-to-vector conversion system is integrated with a modified Gestalt Photomapper(scanner) dedicated primarily to data capture/conversion of analog contour data. Preparation - analog source documents are limited to small format (8" X 8") black and white film positives (or negatives). This usually requires the cutting of an original source into sub-sections. Digitization - a Vidicon camera is mounted in a fixed position on the frame of a Gestalt Photomapper. The source document is affixed to a flatbed transport which moves on fixed scan lines (computer-controlled skipping ranges from fourteen micron to eight inch increments, the later representing a physical, yet unrecommended setting) under the camera. Effective pixel resolution is .00156" X .00156".
Raster-to-Vector Conversion - a proprietary hardware approach which basically applies line thinning and line following logic. This conversion is accomplished for each scanning patch in approximately twenty milliseconds, thus no actual raster data is stored by the system. Some automatic procedures for

vector data review/edit are provided. These include gap closure and stub and snow removal functions. Tagging - no procedures for feature tagging are directly available. Interactive computer graphics procedures for feature tagging can be utilized. Spatial Coding -no spatial coding procedures are directly available with the LINETRAC system. Data Management - some internal file management is maintained.

System Highlights: The reported speed of converting analog contour data to digital vector representation represents the most critical aspect of the system. Its deficiencies include a source document size limitation, and preparation requirements and contour data orientation.

Research and Development: According to David Baker (Director of Applied Technology) and Bob Gogineni (Director of Research and Development), there is some consideration of increasing the size of the scanning mechanism to support greater source input formats. Further work on expanding the intelligent processing of contour data (i.e., automatic feature tagging) is also underway.

DMA Applicability Statement: This is not a commercial system at this time. It has been built to serve in-house requirements. The narrow application to contour data also appears to be a limiting factor. Unless DMA would be interested in exploiting the LINETRAC's hardware R/V convertor for development of a custom contour data capture system, its applicability to overall requirements is quite limited.

6.3.8 Gerber Scientific VDS-2500 Digitizing System

System Summarization: an automatic line-following vector digitizing system adapted to a large format flatbed plotter. It supports analog and digital input, manual/automatic vector digitizing, interactive data editing, file and attribute data management and hardcopy generation.

Preparation - very large source format specifications (5' x 6' or larger) with black and white or single color limitations. Most types of analog input materials are acceptable including paper, Mylar and film positives/negatives. Digitization - a black and white video head (camera) captures 1/2" x 1/2" frames along line segments of analog documents. The operator can also switch to a 4" x 4" camera for larger viewing area. However, the digitizer continues to utilize the 1/2" x 1/2" camera therefore not reducing the resolution. The following digitizing modes are supported: manual, vector advance and automatic. Manual permits operator selection of each coordinate pair defining line segments. Vector advance permits operator intervention of system selected digitization. Automatic allows the system to make all digitizing decisions once a line has been identified and a vector defined. CRT display permits digitized data to be overlaid on the line video (in a choice of colors) to indicate to the operator which data has been captured and which remains uncaptured. Tagging - manual keying-in of up to a ten character (label) identified digitized features is performed via the digitizing control console. Editing - a Graphic Editor is provided to correct, display, modify or change data base elements. Reassignment of tags, or placement of features in one of sixty-four (64) working levels is supported. Spatial Coding - the system is not known to support spatial coding other than standard Gerber (G-Code) data format and "Gerber-out" format for data transfer to Intergraph. Data Management - a sixty-four (64) level file/attribute management facility is provided.

According to Lowell Nerenberg, Sales Engineer for Gerber Scientific Instrument Company, the VDS-2500 is currently (12/84) in its final testing and debugging phases.

Cartographic Data Capture Systems

COMPARE/CONTRAST CHARTS

* Compare/Contrast Charts - Information provided on these charts reflects data provided in industry publications and derived from conversations with representatives of individual systems. Data included in the charts is believed to be accurate and current as of the fall of 1984. However, this does not infer that Battelle has tested these facilities to prove their existence or level of performance.

Symbol Legend

x - function is provided

N/A - not applicable

- not available (blank)

? - no reliable information

I - interactive computer graphics

C - computer-assisted function

A - automatic function (software or hardware)

Figure 7A

PREPARATION

Material Input Requirements

	paper	mylar	scribecote	film pos.	film neg.	pencil	col. pencil	black ink	4-col ballpoint	print col./line	print col. /scrn.	max. # colors	maximum sheet format	maximum readable format	minimum line separation	comments
SCITEX RESPONSE - 280	x	x	x	x	x	x	x	x	x	x	x	12	1016 x 914 mm 40" x 36"	914 x 914 mm 36" x 36"	.002"	ELP scanner/plotter 42" x 75" format
MBB/KONGSBERG SYSSCAN MAP INFORMATION SYSTEM	x	x	x	x	x	x	x	x		x		2-3	600 x 1000 mm 23.6" x 39.3"	600 x 1000 mm 23.6" x 39.3"	.025 mm	
LASERSCAN LASERTRAK					x							B & W	148 x 105 mm 5.8" x 4.1"	98 x 68 mm 3.85" x 2.70"	?	Negative image material input, <u>only</u>
BROOMALL SCAN GRAPHICS ACDS	x	x	?	x	?	x	x	x		x	x	B & W	1526 x 1017 mm 60" x 40"	?	?	
INTERGRAPH/ OPTRONICS SCAN DATA CAPTURE	x	x	?	x	x	x	x	x	x	x	x	3 pass fil-ter	1016 x 1016 mm 40" x 40"	1000 x 1000 mm 39.3" x 39.3"	50 - 75 microns	
ANA TECH VANA	x	x		x	x	x		x					609.6 x 431.8 mm 24" x 17"	scanner dependent	scanner dependent	
TELEDYNE GEOTRONICS LINETRAC				x	x							B & W	9" x 9"	203 x 203 mm 8" x 8"	?	
GERBER SCIENTIFIC VDS 2500 DIGITIZING	x	x	x	x	x	x	x	x		?		B & W	72" x 144"	?	.002"	

Figure 7B

DIGITIZATION

System Specifications/Characteristics

	raster scan	auto-L.-follow	drum	flatbed	tab reg.	max. sheet fmt.	max. readable fmt.	color recog.	color storag	color calib.	grey recog.	grey storage	grey calib.	geometric linearity	stability	point accuracy	point repeatability	resolution	spot size	scan l.p.i.	throughput
SCITEX RESPONSE - 280	X	N/A	X	N/A	X	1016 x 914 mm 40" x 36"	914 x 914 mm 36" x 36"	X	12	X	12 & 64 (ELP)	12 & 12 (ELP)	X	±.0016"	±.0016"	?	?	4-47 lpm & 4-80 (ELP)	25 - 2.5 mm/in.	100-1200 spi 100-2000 (ELP)	18 x 26" x .004" res. x 135rpm = (33 minutes)
MBB/KONGSBERG SYSSCAN MAP INFORMATION SYSTEM	X	N/A	N/A	X	glass or vacum	600 x 1000 mm 23.6" x 39.3"	600 x 1000 mm 23.6" x 39.3"	X	2 - 3	X	256	128	X	±.00039"	±.00039"	±.001" ±25 micr	±.00039" ±10 micr	.001" - .008"	.001"², .002"², .004"², .008"²	125-1000 spi	18 x 26" x .004" res. x 16 mm/s = (7 minutes)
LASERSCAN LASERTRACK	N/A	X	N/A	X	no tab/ neg. hold. only	148 x 105 mm 5.8" x 4.1"	98 x 68 mm 3.85" x 2.7"	N/A	N/A	N/A	N/A	N/A	N/A	±.0004"	?	±.0008"	±.0004"	N/A	.0008"	N/A	T=L/12.5 + F/720 T= time in hours L= line lth meters
BROOMALL SCAN GRAPHICS AGDS	X	N/A	N/A	X	glass	1526 x 1017 mm 60" x 40"	?							±.0016"	±.001"	?	?	.001"	.001"	1000 spi	Data density dependent
INTERGRAPH/ OPTRONICS SCAN DATA CAPTURE SYSTEM	X	N/A	X	N/A	?	1016 x 1016 mm 40" x 40"	1000 x 1000 mm 39.3" x 39.3"	X	3 pass	fil-ters	X	256	X	.002" - .004"	±.0002" - .001"	12.5 microns	± 25 microns	.001" - .008"	.001"	1000 spi	18 x 26" x .004" res. x 1300 rpm = (5 minutes)
ANA TECH VANA SYSTEM	X	N/A	X	X	scan. depen	scanner dependnt	scanner dependnt				X		X	.00125"-.0025" scanner dependnt	scanner dependnt	scanner dependnt	scanner dependnt	scanner dependnt	.00125"-.0025" scanner dependnt	250-1400	Unknown; depend on which scan system used
TELEDYNE GEOTRONICS LINETRAC	X	N/A	N/A	X	neg. hold.	228.6 x 228.6 mm 9" x 9"	203 x 203 mm 8" x 8"							?	?	?	?	?	.00156"	?	Unknown
GERBER SCIENTIFIC VDS-2500 DIGITIZING	N/A	X	N/A	X	?	72" x 144"	?	N/A	N/A	N/A	N/A	N/A	N/A	±.002"	?	±.003"	±.002"	.0005	?	N/A	Unknown

Figure 7C

RASTER-TO-VECTOR CONVERSION

	Raw Raster Data Review/Edit															Raster-to-Vector Conversion					
	run length encoded	entry/exit	other	review only	hard copy	gaps	spikes	stubs	sticks	streaks	snow	coalescence	un-straight lines	un-straight corners	geometric distortion	ease of use	skeletonization	line extract	topology	speed	advanced facilities
SCITEX RESPONSE - 280	x	N/A	x	N/A	x	I C A	I C A	I C A	I C A	I C A	I C A	I C A	I	I		moderate	peel	?	SIF Gerber DIG	?	Some feature recognition (more in development)
MBB/KONGSBERG SYSSCAN MAP INFORMATION SYSTEM	x	N/A	x		x	A	A	A	A	I C	A	I	A	A		moderate	peel medium axis	?	SIF Gerber GSW	?	Pattern recognition feature identification, software development
LASERSCAN LASERTRAK	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	IFF	N/A	Digitizing output is vector; no R/V conversion required as post-process
BROOMALL SCAN GRAPHICS AGDS		x		x												simple	?	?	?	?	Realtime rasterization/vectorization software
INTERGRAPH/ OPTRONICS SCAN DATA CAPTURE	x	N/A	data compression formats	x	x						A					simple	N/A	scanline	L/N Poly SIF IGES	?	Auto-feature tagging Auto-character recognition
ANA TECH VANA																?	N/A	hardware	?	?	Proprietary real-time hardware R/V conversion process
TELEDYNE GEOTRONICS LINETRAC						A	A	A			A	A				?	peel	line follow	?	?	Patent pending on hardware implementation of R/V conversion
GERBER SCIENTIFIC VDS-2500 DIGITIZING	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Gerber	N/A	Digitizing output is vector; no R/V conversion required as post-process

Figure 7D

RASTER-TO-VECTOR CONVERSION

Raw Vector Data Review Edit

	gaps	spikes	stubs	sticks	streaks	slivers	coincident points	coincident lines	snow	unsquare corners	unstraight lines	displaced nodes	skewed intersections	wandering centerlines	topological/spatial	ease of use	comments
SCITEX RESPONSE - 280	I C A	I C A	I C A	I C A	I C	I A	I A	I A	I A	I A	I A	I A	I A	I	?	moderate	Vector display over raster image
MBB/KONGSBERG SYSSCAN MAP INFORMATION SYSTEM	I A	I A	I A	I A	I	I A	A	A	I A	I A	I A	I	I		?	moderate	
LASERSCAN LASERTRAK	I	I	I	I	I	I			I	I C	I C	I	I	I	?	moderate	
BROOMALL SCAN GRAPHICS ACDS	I	I A	I A	I A	I	I			I A	I	I	I	I	I		moderate	
INTERGRAPH/ OPTRONICS SCAN DATA CAPTURE	I C A	I C A	I C A	I C A	I	I			I C A	I	I A	I C	I C	I	? C A	moderate	Vector display over raster image
ANA TECH VANA	I A	I A	I A	I A	I A	I A	I A	I A	I A	I A	I A	I A	I A	I A		?	
TELEDYNE GEOTRONICS LINETRAC	A	A	A	A					A							automatic	
GERBER SCIENTIFIC VDS-2500 DIGITIZING	I	I	I	I	I	I			I	I	I	I	I	I		?	

Figure 7E

	TAGGING								SPATIAL CODING			
	contours	intersecting networks	merging networks	polygons	point symbols	symbolized elements	labelling	ease of use	universal referencing	topological encoding	data structuring	comments
SCITEX RESPONSE - 280	I C A	I	I	I C A	I C A			moderate	2-3 projection transformation routines	node lists link/node structure	digits SIF Gerber, etc.	
MBB/KONGSBERG SYSSCAN MAP INFORMATION SYSTEM	I C	I	I	I C A	I			moderate	x	DLG under development	GINNIS SIF Applicon IBM-CADAM	
LASERSCAN LASERTRAK	I	I	I	I	I			moderate	some transformation routines available	link/node structure	IFF	
BROOMALL SCAN GRAPHICS ACDS	I	I	I	I	I			moderate	UTM-lat/long	link/node structure DEM	ACDS DLMS	
INTERGRAPH/ OPTRONICS SCAN DATA CAPTURE	I C	I C	I C	I C A	I A			moderate	approximate. twenty trans. routines available	under review	SIF IGES	
ANA TECH VANA												Automatic lineweight, character and symbol recognition is supported.
TELEDYNE GEOTRONICS LINETRAC												No tagging or spatial coding facilities are currently provided
GERBER SCIENTIFIC VDS-2500 DIGITIZING	I	I	I	I	I	I		?		link/node	Gerber	

7.0 A DMA STANDARD CARTOGRAPHIC BENCHMARK TESTING CAPABILITY

7.1 A Definition of Benchmark Testing

Benchmark testing is defined as a set of procedures which are implemented to measure and verify the level of performance of a machine or system in the context of objective criteria.

7.2 A Standard Benchmark Capability for DMA

A standard benchmark testing capability provides the Defense Mapping Agency with an objective means of evaluating state-of-the-art automated cartographic data capture systems. It establishes a sound basis for assessing the potential applicability of individual systems to DMA's digital cartographic data/map production requirements. It also provides a rational framework for system comparison.

Battelle has designed a standard cartographic benchmark testing methodology based upon comprehensive evaluation of the DMA analog-to-vector conversion process, individual map/data product specifications and production requirements, and a topical investigation of state-of-the-art data capture systems and technology. The two major components of this benchmark testing capability include: a standard materials testing package and a standard set of testing procedures.

The technical scope of the benchmark testing capability is limited to basic analog-to-vector conversion processing technology. Specifically, it addresses the procedures and problems associated with the capture and conversion of elemental cartographic geometries found on standard DMA analog source manuscripts. These geometries consist of point, line and area cartographic elements possessing limited symbology (i.e., line thickness variation and limited dashing, only). This does not imply an ignorance or denial of the emerging importance of advanced "intelligent" A/V systems and techniques.[13] It does emphasize, however, the fundamental role and inherent complexities of basic A/V processing and also reflects the initial intent of this research project.

7.3 Standard Benchmark Testing Materials*

Two basic types of benchmark materials are included in the testing package: standard DMA products and synthetic testing materials. Samples of standard DMA products are chosen from Digital Feature Analysis Data (DFAD) photogrammetric compilation manuscripts, Digital Terrain Elevation Data (DTED) cartographic overlays, and Hydrographic Chart compilation manuscripts. All three are included for the following reasons: 1) They represent the three basic analog inputs to the current DMA digital cartographic data/map production program. 2) Combined, they contain many of the typical cartographic geometries and anomalies. 3) All data capture systems must prove themselves capable of processing typical DMA analog source manuscripts to a minimum standard.

7.3.1 Standard DMA Products

The DFAD compilation manuscript is a good composite example of point, line and polygonal cartographic geometries, represented by hand-drawn color pencil on a Mylar base. The sample DFAD submitted to the benchmark package should be an original or a hand-drawn copy, with variable data coverage throughout. This testing material will provide information about a data capture system's ability to handle four different phenomena: typical DFAD geometry, typical DFAD data density, color recognition, and geometric errors or anomalies found on a DFAD sheet.

[13]This is discussed in Section 8.0, "Recommendations for Future Research and Development".

* Copies of benchmark testing materials are provided in a separate envelope accompanying this report.

The DTED contour and drain/ridge overlays are good examples of non-intersecting isolines and merging networks, respectively. The contour overlay sample should have various levels of data coverage throughout the sheet, representing different types of terrain. This should naturally result in a companion drain/ridge overlay of similar complexities and coverages. Depression contours represented by dashed lines and index contours of heavier lineweights should also be evident on the sheet. These testing materials will provide information about a data capture system's ability to process five different phenomena: typical DTED geometries, representative data densities, lineweight variations, minor symbolization and geometric errors or anomalies found on DTED contour and drain/ridge overlays.

The Hydrographic Chart compilation manuscript sample is a good example of different cartographic geometries, hand-drawn in multiple colors on a Mylar base. The sample chosen should be representative of the variety of data types, coverages and typical compilation colors normally found on such a manuscript. This sample material will provide information about a data capture system's facility in processing data possessing the following characteristics: typical Hydrographic Chart geometries, typical data densities, multiple colors, large format sheet size, and geometric errors or anomalies often found on Hydrographic Chart compilation manuscripts.

7.3.2 Synthetic Testing Materials

The rationale for creating synthetic testing materials is based on a requirement to test the performance levels of state-of-the-art data capture systems in processing critical aspects of DMA analog source manuscripts. These materials also afford an opportunity to perform such testing under controlled circumstances. The synthetic testing materials, synthetic test sheet #1 and synthetic test group #2, both focus on three particular characteristics: unique cartographic geometries, geometric degradation, and data density ranges.

7.3.2.1 Synthetic Test Sheet #1

This test sheet portrays a series of "perfect" cartographic geometries in the first column. These represent the basic cartographic geometric characteristics described earlier in the report. They include: non-intersecting lines (e.g. concentric circles), merging networks (e.g. "Y's"), intersecting networks (e.g. +++), T intersections (e.g. T), and limited symbologies (e.g. --- - -). Error/anomaly types for each geometry are arrayed in their respective X axes. Examples of these include gaps and spikes of controlled dimensions (i.e., in the thousandths of inches). Additionally, each row of unique geometries is repeated two times in different lineweights. The resulting three lineweights represent the range of typical lineweights found on DMA analog source manuscripts (i.e., .004", .008", .012"). The actual test sheet consists of solid black lines on a clear film base.

Synthetic test sheet #1 provides a controlled medium to evaluate different systems capabilities in processing the variety of cartographic geometries found on typical DMA products. It facilitates the testing of error detection/correction routines provided by most state-of-the-art systems. It also permits a study of errors introduced by systems during data processing, particularly those associated with vectorization and error correction routines.

7.3.2.2 Synthetic Test Group #2

Synthetic test group #2 consists of twelve separate input manuscripts. These subdivide into three types of cartographic geometries (concentric circles, merging "stream" networks and orthogonal grids) each plotted in four levels of increasing density. These test patterns are portrayed as solid black lines on a clear film base. They are all plotted at identical lineweights with known lengths of linear inches.

The primary utility of test group #2 is to support analysis of the impact of geometric formation and data density on raster-to-vector conversion throughput. The basic question is whether algorithms for R/V conversion are written for specific geometries (e.g., contours) and whether increasing data density results in a linear function for throughput performance. A by-product of testing these materials would be a rating system for different systems indicating their strengths and weaknesses in processing certain cartographic input. For example, some systems may perform well in converting contours to vector center-lines but have difficulty in converting road networks. The converse may be true for other systems.

7.4 Benchmark Testing Procedures

Benchmark Testing Procedures fall into several basic categories: raster scanning (or alternative digitization methods), automatic raster editing, raster-to-vector conversion, automatic vector editing, and plotting on an automatic vector plotter. Evaluation criteria are based on individual process timings, combined process timings, virtual image quality assessment, digital plot/analog input "overlay" analysis, system integration/user friendliness evaluation, and statistical tests.

It should be noted that the benchmark testing does not include analog source preparation, feature tagging, interactive editing or spatial coding procedures. This is due to time limitations placed on the testing at DMAHTC and the difficulty of properly judging the more subjective (human oriented) procedures. Although such human factors testing is possible it was deemed infeasible for this particular application at DMA.

7.4.1 Process Timings

Time is a basic criteria for performance evaluation. Recording of execution times is critical for each procedure in the analog-to-vector conversion process delineated in section 7.4. This data provides a gauge to evaluate procedures implemented with different types and densities of data. It also permits system comparisons.

7.4.2 Combined Process Times

This refers to the aggregation of timing statistics for specified processes. For example, certain systems implement raster-to-vector conversion in multiple steps while others perform R/V conversion in a single observable process. Summation of multiple step process times facilitates comparison with times for "one-step" procedures.

7.4.3 Virtual Image Quality Assessment

This refers to a visual evaluation of geometric image quality on a virtual ("CRT") device. Samples of cartographic geometries should be observed for marked changes or quality degradations as data is processed through a system. This is particularly addressed to processing Synthetic Test Sheet #1. Perfect geometries should be evaluated throughout raster-to-vector conversion and repetitive automatic editing phases. Unique permutations and errors should be graphically noted along with pertinent information (e.g., error type "a" occurred after xxx procedure was implemented).

7.4.4 Digital Plot/Analog Input "Overlay" Analysis

The overlaying of film positive digital plots (of processed data) with original test material input documents supports the following evaluation:

a) centerline accuracy assessment; how well does the vectorized data align with the center of original analog lines?

b) geometric degradation; what kinds of errors and anomalies have been generated (or eliminated) by the cumulative raster-to-vector conversion procedures?

7.4.5 System Integration/User Friendliness Evaluation

This reflects a somewhat subjective assessment of the integration of hardware, software and procedural components of an automated cartographic data capture system. This assessment extends to an evaluation of how accessible these components are to a system operator, and how well data is processed by linking one routine (or process) to another.

7.4.6 Statistical Tests

Statistical analysis can be applied to timing data to establish certain procedural relationships between types of cartographic data and A/V conversion functions. For example, regression analysis can be applied to evaluate the relationships between raster scanning time and size/density of input analog manuscripts. Based on repetitive sampling it is possible to predict the dependent variable (T - Time) as a function of manuscript sheet size, data density, scan resolution, and an error parameter. Similarly, the impact of increasing data density on R/V conversion times can also be predicted by statistical analysis. Benchmark testing results should be evaluated using these statistical techniques where deemed valid.

7.5 Testing Standard DMA Products

Each sample DMA standard product will be processed in the following manner: 1) scan manuscript and note scan time 2) review raster image to assure successful scan 3) run raster-to-vector conversion routines and time each individual component of the process (e.g., thin time, vectorization time) 4) review vector image to assure successful conversion 5) create digital plot on vector plotter (film positive is preferred).

The DFAD manuscript should be processed according to the above guidelines with the following modification: all feature tags should be automatically removed from the file either during or directly after the scanning phase, and certainly before raster-to-vector conversion (unless a system claims no time impediment from converting tags and other annotation).

The DTED contour and drain/ridge overlays should be processed separately and each according to the above guidelines.

The Hydrographic Chart manuscript should be processed according to the above guidelines with the following modification: all bathymetric soundings should be automatically removed from the file either during or directly after the scanning phase, and certainly before raster-to-vector conversion (unless a system claims no time impediment from converting tags).

The central objective of submitting samples of DMA standard analog cartographic source manuscripts to the benchmark test is to assess the basic performance level of state-of-the-art data capture systems in assimilating these typical inputs. The keeping of timing statistics and the completion of an overlay analysis (for center-line accuracy and error detection/correction analysis and analysis of error propagation) establish two performance evaluation criteria.

7.6 Testing Synthetic Benchmark Materials

This section delineates the procedures for benchmark testing Synthetic Test Group #2 and Synthetic Test Sheet #1.

7.6.1 Synthetic Test Group #2

*Two basic testing procedures are recommended for the synthetic benchmark materials. Synthetic test group #2 should be processed in the following manner: 1) scan manuscript and note scan time 2) review raster image to assure successful scan 3) run raster-to-vector conversion routines and time each component of the process (e.g., thin time, vectorization time) 4) review vector image to confirm successful conversion 5) create digital plot on vector plotter (film plot with lineweight thinner than original preferred) 6) overlay digital plot with analog input to perform accuracy assessment (center-line representation) and note error/anomaly generation. This process applies equally to all twelve input sheets comprising test group #2.

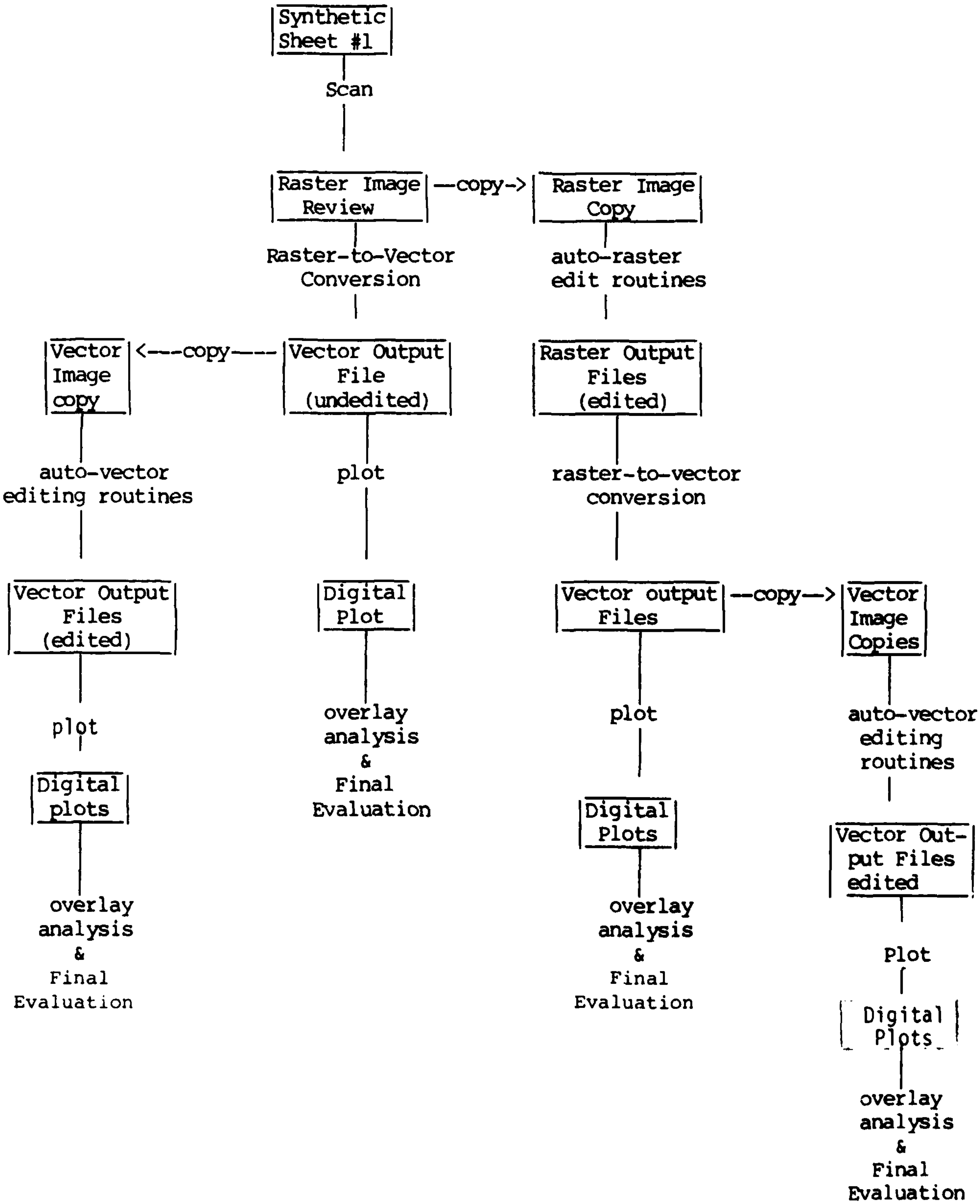
7.6.2 Synthetic Test Sheet #1

Four parallel benchmark testing procedures are recommended for Synthetic Test Sheet #1 (refer to Figure 8A for a graphic overview). The first procedure includes the following five steps: 1) scan input manuscript and record scan time 2) review raster image to confirm successful scan 3) run raster-to-vector conversion routines and record time for each component of the process (e.g., thin time, vectorization time) 4) create digital plot (on vector plot with .003" lineweight) 5) perform overlay analysis with original analog input manuscript for accuracy assessment (center-line representation) and record error/anomaly generation.

The second procedure includes the following five steps: 1) make copy of raster image resulting from scan in the first procedure 2) run and store results of successive raster editing routines (e.g., run a series of gap closure routines with changed gap tolerances, run a series of spike removal routines with changed spike tolerances, run a series of snow removal routines with changed pixel cluster size tolerances, etc.) 3) run raster to vector conversion routines on individual raster files and record time for each component of the process 4) create individual digital plots (on vector plotter with .003" lineweight) 5) perform overlay analysis with original analog input manuscript for accuracy assessment (center-line representation), error/anomaly correction performance evaluation and recording of error/anomaly generation.

Figure 8A

BENCHMARK PROCEDURES FOR SYNTHETIC TEST SHEET #1



The third procedure includes the following four steps: 1) make copy of a selection of vector output files generated in the second procedure 2) run and store results of successive automatic vector editing routines (e.g., gap closure with changed tolerances, spike removal with changed tolerances, snow removal with changed tolerances, etc.) 3) create individual digital plots (on vector plotter with .003" lineweight) 4) perform overlay analysis with original analog input manuscript for accuracy assessment, error/anomaly correction performance evaluation and recording of error/anomaly generation. Overlay analysis between digital plots from second and procedures is also recommended.

The fourth procedure includes the following four steps: 1) make copy of unedited vector output file from the first procedure 2) run and store result of successive automatic vector editing routines (same routines as in previous procedures) 3) create individual digital plots (on vector plotter with .003" lineweight) 4) perform overlay analyses with original analog input manuscript for accuracy assessment, error/anomaly correction performance evaluation and recording of error/anomaly generation.

The first procedure offers an opportunity to evaluate a data capture system's capability to perform analog-to-vector conversion of a representative sample of DMA cartographic geometries. Errors or anomalies input by a system during this processing will be readily identifiable. The second procedure focuses on systems with automatic raster editing capabilities. Evaluation of the performance level of these editing routines is thus supported. The third procedure is a direct follow-on of the previous one. In this case a system with "parallel" automatic raster/vector data editing facilities is tested. The effectiveness of the vector editing procedures is evaluated. Overlay analysis of the result of raster editing (only) and raster/vector editing (combined) highlights the value of "redundant" processing as well as reveals new problems which might be associated with such a process. The fourth procedure focuses on those systems with data editing routines solely (or mainly) in vector mode. Here again their effectiveness can be evaluated against a standard analog benchmark material and results compared to systems offering similar or alternative editing schemes.

8.0 RECOMMENDATIONS FOR FUTURE RESEARCH AND DEVELOPMENT

There are many paths DMA might choose to follow in future research and development of the analog-to-vector (A/V) conversion process. Battelle recommends six (6) specific areas for future investigation based upon our work during the past year. These represent applied and basic research concepts which address critical aspects of the A/V conversion process.

8.1 Advanced Benchmark Capability

The analog-to-vector conversion benchmark testing package and methodology presented in this report represents a basic capability. It specifically addresses the methods and problems associated with the conversion of minimally symbolized point, line and area analog cartographic features. Recently, significant progress has been reported by several commercial vendors in the areas of automatic pattern/symbol/character recognition, automatic feature tagging/contour elevating and computer-assisted spatial coding procedures. In light of these advancements in automated cartographic data capture technology, the development of an advanced benchmark testing package and methodology is highly recommended.

The envisioned advanced benchmark testing package would build upon the basic capability already developed. It would integrate the broad spectrum of analog cartographic source materials including complex geometries and symbolization. Specific tests would be developed to evaluate advanced data processing techniques described above.

8.2 Optical Image Processing for Cartographic Data Capture

This proposed research would focus on the feasibility of developing an intelligent feature recognition optical raster scanning capability for cartographic data capture. Such analysis would be predicated on state-of-the-art optical image correlation methodologies, some of which have been developed by the Battelle Columbus Laboratories Optics Department. This approach would significantly differ from other techniques for automatic feature recognition which apply template matching and other statistical tests to raster data images via software. The potential speed improvements alone warrant a serious look at optics technology in this arena.

8.3 Vendor System's Optimization and Upgrade

Battelle recommends that DMA initiate a program to analyze, catalogue, and optimize the use of all software routines currently available on in-house automated cartographic data capture systems. This should be done within the framework of DMA analog-to-vector conversion requirements with an eye towards taking full advantage of all system components. Battelle further recommends that a prototype program commence with the Scitex system currently in use at DMAHTC. A significant number of functions exist which do not have clear applications to cartographic processing. In other cases, multiple functions perform similar tasks with different results, depending upon data and particular applications. Battelle firmly believes that more effective use of automatic functions, batch programming and heretofore unused options is possible. A functional system of ongoing software evaluation for optimal use will support this goal.

8.4 Advanced Computer Architectures for Cartographic Data Processing

Battelle recommends that DMA support future research and development of advanced computer architectures for cartographic data processing. These might include array processors, multiple CPU's, VLSI, parallel processing for raster-to-vector conversion, feature recognition and automatic tagging. DMA would benefit from this ongoing research by successful system developments and maintenance of in-house understanding of state-of-the-art computer architectures. The latter is particularly important for informed evaluation of commercial automated cartographic data capture systems.

8.5 Basic Research of Raster-to-Vector Conversion Algorithms

Battelle recommends that DMA support continued research into the development of new raster-to-vector conversion algorithms. These developments may reflect work done in related fields (e.g., image processing, electrical engineering, computer graphics) or new processing environments. The latter may be relevant to research in advanced computer architectures. Emphasis should be placed not only on throughput efficiencies but on more sophisticated and effective methods of handling data anomalies. Battelle's evaluation of, and experience with, modern automated cartographic data capture systems has revealed that many data editing routines fail to detect/correct a large percentage of errors with current algorithms. This requires a continuation of labor intensive, time consuming "clean-up" procedures.

8.6 Data Structures for Cartographic Data Processing and Application

It is recommended that DMA support future research and development of geographic/cartographic data structures as they relate to the A/V conversion process. The optimal use of vector and raster data structures and the utility of hybrid structures such as "vaster"* should be evaluated. Research should address questions about data structure influence over editing functions, automatic symbol/feature recognition and long term revision methodologies.

* A term coined by professor Donna Peuguet which refers to an integration of vector and raster characteristics into a single data structure.

Appendix A

DMA ANALOG-TO-VECTOR MODEL GLOSSARY

DMA ANALOG-TO-VECTOR MODEL GLOSSARY

PREPARATION - refers to the preparation of analog cartographic materials for conversion to digital representation. Two types of materials are developed during this phase: analog graphic and analog feature attribute. Analog graphic is any cartographic "paper" product containing linear (point, line, area) elements, derived from either photogrammetric compilation or printed manuscript. The preparation of analog (carto)graphic manuscripts for digital conversion occurs in four stages.

Analog Graphic

1. Data Selection - the process where data categories are selected from a cartographic manuscript to be converted to digital representation in the digitization phase. In some cases, all analog cartographic data is required, whereas under different conditions only select data categories may be chosen for digital conversion.

   At the Defense Mapping Agency (DMA) all data on the photogrammetric compilation of Digital Feature Analysis Data (DFAD) is submitted to the digitization phase. Alternatively, only the contours, drainage lines, and ridge lines are required for a Digital Terrain Elevation Data (DTED) cell, thus omitting the road network on the original topographic manuscript, among other data categories.

2. Manuscript Reformat - the transformation of analog cartographic manuscripts prior to digitization. In some cases, analog manuscripts may require recompilation, photographic reduction, panelling or subdivision into multiple pieces due to size or image density. In all events, manuscript reformatting reflects an initial incompatibility between the input cartographic document and a particular digitization system.

   At DMA, DTED contour relief manuscripts are sometimes recompiled, retraced or are linework enhanced and photographically transformed several times (e.g., film negative-film positive) before final submission to the digitization phase.

3. Manuscript Enhancement - the enhancement of analog cartographic manuscripts to provide additional data coverage, assist in topological encoding, and generally improve the efficiency of the analog-to-vector conversion (A/V) process. The availability of color raster scanners provides improved capabilities which are better utilized as a result of prior manuscript enhancement. Node detection can be facilitated by unique color marks on the input manuscript, at bridge intersections for example. Feature separation can be accomplished by outlining unique feature types with different colors on a single compilation overlay.

An example of manuscript enhancement at DMA is found in the creation of an additional overlay containing ridge and drain lines, supplemental contours, and hachures in the production of DTED cells. This is done to provide greater accuracy in the resulting elevation matrix.

4. Manuscript Review Edit - the review and edit of analog cartographic manuscripts prior to digitization to assure data quality, completeness, and efficient utilization of resources. Analog materials should be "clean" and absent of unwanted graphic images. Gaps in "continuous" lines and merging of "discontinuous" lines requires editing, either at this stage or later in the analog-to-vector conversion process. Analog manuscripts containing sparse graphic coverage may be directed to manual digitization instead of raster scanning, in a more efficient utilization of available resources.

Analog feature attribute preparation refers to the manual process of preparing feature code overlays which annotate codes for each type of cartographic element to be digitized and processed. This analog feature attribute data is utilized during the tagging phase. Another type of feature attribute preparation is performed by creating separate cartographic feature overlays, or multi-color cartographic proofs (i.e., each feature type is represented by a unique color), or overlays where each "unique" cartographic feature type is in a different lineweight, prior to digitization. Depending on the system of digitization, some limited form of automatic feature tagging may be facilitated by this preparation. As more sophisticated methods of feature tagging are developed it is expected that resources dedicated to manual feature attribute preparation and feature tagging will be reduced dramatically.

At DMA analog feature attribute preparation is a critical and time consuming activity. Both DTED and DFAD require the assignment of codes which indicate elevation and/or feature type information. These codes are then utilized during the tagging operation (performed on the AGDS tag/edit and Lineal Input System (LIS) systems). The Scitex at DMAHTC provides for the "automatic" feature separation of analog cartographic manuscripts which facilitates the efficient tagging of digital cartographic data.

DIGITIZATION - refers to the conversion of analog cartographic materials to digital computer representation by the individual or combined means of electronic, optical, laser, solid state, and mechanical processes. The three most common forms of digitization are manual, automatic line-following, and raster scanning.

1. Manual Digitization - the "manual" process where point, line and areal cartographic elements are located via a hand-held cursor, identifying the x,y coordinates of the elements on an electronic mesh attached to the digitizing table. The x,y "table" coordinates are stored on a magnetic tape or disk storage system. Output of this manual digitization is described as vector cartographic data. Alternative methods of manual digitization include line tracking and error correcting systems.

Currently, both the AGDS and ACDDS systems at DMA contain manual digitizing facilities of the type first described above.

2. Automatic Line-following Digitization - the "automatic" process where point, line and areal cartographic elements are located via an automatic line-following cursor, possibly a deflected laser beam, identifying their x,y coordinates. This is usually an interactive (i.e., human operator/computer-assisted) process where some logical decisions are automatically made based on predetermined parameters and at other times human intervention is required for decision-making. For example, the direction to continue line-tracking once a cursor has encountered a network intersection can be determined by the system or by the human operator. The coordinates captured by the digitizer are stored on a magnetic tape or disk system. Output of this digitizing process is described as vector cartographic data.

DMA does not own or operate an automatic line-following digitization system at the present time (1984).

3. Raster Scanning Digitization - the automatic process where point, line, and areal cartographic elements are converted to digital representation by means of electro-optical, solid state, or

laser "cameras" which traverse the manuscript (mounted either on a rotating drum or stationary flatbed) in scan lines at prescribed resolutions (some systems have fixed resolutions and others provide for calibration). Certain scanners are color sensitive (storing up to twelve colors in one pass) and others are strictly black and white scanners (detecting and storing the transitions from black to white at a predetermined threshold). Data is stored in a raster, two-dimensional data structure on magnetic tape or disk storage systems. Alternative data compaction schemes include run length encoded and entry/exit point data formats. This raster data often requires a raster-to-vector conversion before the data can be fully exploited by "traditional" digital cartographic systems.

Stored Raster Data - cartographic data which is stored in a raster format usually for the intended purpose of eventually producing a digital plot.

Linear cartographic features are scanned on the Scitex raster scanner, converted to a center-line representation and then reconverted to a raster format fully symbolized. This data can then be plotted as a fully symbolized cartographic feature plot on a raster plotter.

It is also conceivable that raster data collected by agencies external to DMA, may be fed into the digital cartographic production system as stored raster data. This data can either remain stored in this format or be converted to vector data and continue processing in a "standard" manner.

RASTER-TO-VECTOR CONVERSION - refers to the conversion of digital cartographic data from a raster data structure to a x,y vector, centerline representation. In addition to the central raster-to-vector conversion process, there are pre-processing steps (i.e., raw raster data review/edit) and post-processing steps (i.e., raw vector data review/edit).

1. Raw Raster Data Review Edit - the process where raw raster data (i.e., direct output of raster scan digitizing) is reviewed for data quality, completeness, and anomalies (e.g., gaps, stubs, spikes, sticks, streaks, coalescence, misalignment) commonly associated with the digitization process (although often attributable to the input cartographic manuscript). This process may be accomplished interactively (i.e., through interactive computer graphics techniques), via computer-assistance (i.e., software functions identify "problems" and facilitates their correction through interactive computer graphics techniques) or automatically (i.e., software/hardware functions detect and correct errors or anomalies in the data).

At DMA, all three approaches to raw raster data review/edit are utilized on the Scitex data capture system. The AGDS does not provide for any raster editing functions other than review on a black and white CRT.

2. Raster-to-Vector Conversion - the actual conversion of raster cartographic data to centerline vector cartographic data by means of software and/or hardware configurations. Raster-to-vector conversion typically follows three steps: skeletonization (or line thinning) - reduction of raster elements to one unit of resolution, located in the geometric center of the cartographic feature, line extraction - the identification of unique line segments, and topology reconstruction - the explicit identification of the spatial relationships of points, lines and areas in the vector file. Recently, new approaches to raster-to-vector conversion have been identified. In these cases, hardware processors derive boundary chain vectors from edges of raster data concurrent with the scanning process. This is followed by software determination of vector center-line data. Lineweight tagging results as a by-product of this process.

Many algorithmic solutions to the R/V conversion process have been developed by commercial vendors, academic researchers and government agencies. Their utility and efficiency has been variously dependent upon data type (i.e., non-intersecting lines, closed polygons, merging/intersecting networks, crossing/intersecting networks and point features), cartographic symbolization, cartographic labeling, data volume and geometric complexity. Another important issue is the utility of the resultant vector data format for expected applications.

At DMA, raster-to-vector conversion is implemented on the AGDS vectorization subsystem and the Scitex editing station.

Raw Vector Data Review/Edit - the process where "raw" vector data (i.e., direct output of manual and automatic line following digitization or the raster-to-vector conversion process) is reviewed for data quality, completeness, and anomalies (e.g., gaps, stubs, spikes, sticks, streaks, coalescence, misalignment) commonly associated with digitization and R/V conversion (although can be a "carry-over" from the original cartographic manuscript). This process may be accomplished interactively via computer graphics techniques, by computer-assistance where software functions identify "problems" and facilitate their correction through interactive computer graphics methods, or automatically via software/hardware functions which detect and correct errors and anomalies in the data.

At DMA, raw vector data review/edit is currently performed (by all three methods) on the AGDS vectorization and edit/tag subsystems in addition to the Scitex edit station.

TAGGING - refers to the process of logically and physically associating vector cartographic data with analog feature attribute information in a digital computer file.

1. Analog Feature Attribute/Vector Data Merge - digital vector data is linked with analog attribute information (codes) by means of interactive graphics techniques, software computer-assistance, or fully automated methods.

At DMA a distinction is made between tagging and elevating Elevating is the digital linking of an elevation value to a relief element (i.e., contours, ridge lines, drainage lines). Tagging is the digital linking of a feature code to a vector data element. DMA is supporting the experimental development of Automatic Feature

Tracking software which performs vectorization and feature tagging on raster cartographic data. Potentially fully automated methods of tagging may result from the development of sophisticated pattern recognition/feature extraction software. Tagging at DMA is currently accomplished on the AGDS edit/tag subsystem and the ACDDS. Some software assisted contour tagging/elevating routines are available for the Scitex at DMA as well.

2. Tagged Vector Data Review/ Edit - the process of assuring the correctness and completeness of the tagging procedure. In addition to in-process quality assurance and graphic plotting, interactive computer graphics techniques and fully automated procedures can be implemented for tagged vector data review/edit.

Currently at DMA, vector plots of tagged feature data are made on automatic drafting machines. The plots are reviewed for tagging completeness and accuracy. Corrections are made to the tagged base file. Additional plots are made until all corrections are completed.

SPATIAL CODING - refers to the logical and physical definition of the cartographic data file's internal and external spatial relationships. Three types of spatial relationships are defined: universal referencing, topological encoding, and data structuring. A formal quality assurance and edit function is required to maintain the spatial integrity of the data files.

1. *Universal Referencing* - the conversion of a cartographic data file stored in table (Cartesian) coordinates to a universal reference system (e.g., Universal Transverse Mercator - UTM, Latitude/Longitude, Lambert Conic Conformal).

   At DMA for example, the DTED contour/relief data files are transformed from (table) UTM to geographics during the post-processing phase on the UNIVAC computer.

2. Topological Encoding - the explicit definition of the internal spatial relationships of a digital cartographic file. Such topological encoding can result in simple link/node lists or sophisticated data organizations (based on principles of mathematical/geometric topology) which define the point, line, area relationships within the data file.

At DMA the DTED matrix data structure has an implicit topological identification system where every z-value and its neighbor can be located by row/column positions. The Digital Land Mass Blanking System supports another type of topological data structure.

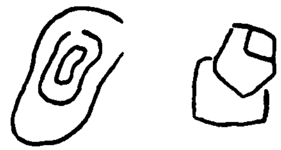
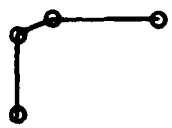
3. Data Structuring - the explicit definition of the data file format. The number of fields, their content and form are all defined in this activity.

The Defense Mapping Agency supports two standard data structures; DLMS and the Standard Lineal Format (DMASLF).

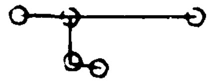
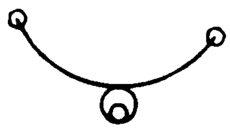
DATA MANAGEMENT - refers to the processing, storage, manipulation and retrieval of digital cartographic vector data.

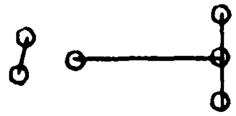
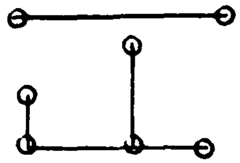
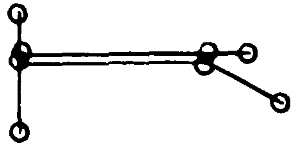
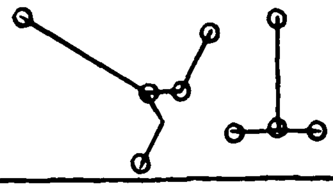
Appendix B

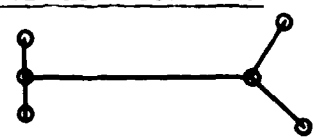
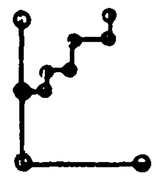
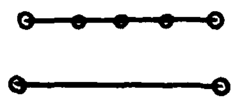
ERROR/TYPE CHARTS

ERROR TYPE	ERROR DESCRIPTION	GRAPHIC ULATION	ERROR CAUSE	ERROR CORRECTION
Gaps	*Breaks in digital or compiled line segments		Analog manuscript Raster scanning R/V conversion Manual digitizing	Manual Computer graphics Auto-software
Coalescing Contours	The connection of at least two contour lines		Analog manuscript Raster scanning R/V conversion Manual digitizing	Re-scribing Photo-spreading Computer graphics Auto-software
Unsquare Corners	Corners whose line segments do not meet at a 90 degree angle		Raster scanning R/V conversion Manual digitizing	Computer graphics Auto-software
Unstraight Lines	Line segments whose direction is not constant		Raster scanning R/V conversion Manual digitizing	Computer graphics Auto-software
Non-skeletonized Line Vectors	Line vectors of greater than one unit of resolution		R/V conversion	Computer graphics Auto-software

*Definitions furnished by H. Jackson--DMAAC

ERROR TYPE	ERROR DESCRIPTION	GRAPHIC SIMULATION	ERROR CAUSE	ERROR CORRECTION
Straight Stub	*A short, two coordinate node segment that is attached at one or both ends to other node segments		Analog manuscript Thinning of variable width raster data elements	Interactive computer Computer-assisted Auto-software
Hooked Stub	A short, (n) coordinate node segment(s) that is attached at one or both ends to other node segments		Analog manuscript Thinning of variable width raster data elements	Interactive computer Computer-assisted Auto-software (?)
V Stub	A short, three coordinate node segment which is attached at both ends to the same node segment		Thinning of variable width raster data elements	Interactive computer
Circular Stub	A circular node segment attached to another node segment		*Thinning* of variable width raster data elements, particularly on curved elements	Interactive computer
Donut Stub	A circular node segment within a larger circular node segment attached to another node segment		Thinning of variable width raster data elements, particularly on curved elements	Interactive computer
Tube Stub	An elongated, tubular node segment attached at both ends to the same node segment		Thinning of variable width raster data elements	Interactive computer

ERROR TYPE	ERROR DESCRIPTION	GRAPHIC SIMULATION	ERROR CAUSE	ERROR CORRECTION
Sticks	*A short, two coordinate node segment not attached to any other segment in the file		Analog manuscript	Interactive computer Computer-assisted Auto-software
Streaks	*A long, two coordinate node segment which may or may not be attached to another line segment		Raster scanning	Interactive computer
Slivers	*Two features which share a common node segment; when thinned feature formats retain different coordinate points resulting in overlap and underlap between features		Manual digitizing Analog manuscript Raster scanning	Interactive computer Auto-software
Coincident Point	*Duplicate sequential points in a line string		Vectorization	Auto-software
V Intersection	Intersection incorrectly defined after raster-to-vector conversion		Manual digitizing R/V conversion	Interactive computer Auto-software
Corner Shortcut	Failure to correctly define the centerline of a curve		Thinning Manual digitizing	Interactive computer Auto-software

ERROR TYPE	ERROR DESCRIPTION	GRAPHIC SIMULATION	ERROR CAUSE	ERROR CORRECTION
Coincident Line	*A duplicate line segment with only two points which have identical start and end nodes as another two point line segment		Manual digitizing R/V conversion	Auto-software
Duplicate Point	*Non-sequential identical coordinate point values in a line string		Manual digitizing R/V conversion	Auto-software
High Power Noise	*Spikes or loops in digital data defined by large directional change with slight distance change		R/V conversion	Computer graphics Auto-software
Low Power Noise	*Line segment which contains more coordinate points along a straight line than are necessary to depict the segment. Slight variation from straight line direction		Manual digitizing Raster scanning R/V conversion	Auto-software
Noise	*Random disturbance in digital data that obscures the clarity or quality of the data		Analog manuscript Manual digitizing Raster scanning R/V conversion Random "glitches"	Manual Computer graphics Auto-software
Wandering Center-line	Vectorization fails to correctly define and maintain the center-line of raster data		R/V conversion	Computer graphics Auto-software

Appendix C

"LITERATURE" SEARCH LISTS

CARTOGRAPHIC DATA AQUISITION SYSTEM

Allam, Dr. M. Mosaad, "Acquisition of Digital Topographic Data and the Need for a Standardized Digital Data Base,": ISPRS Commission IV Symposium, 1982, 1-11.

ANA Tech Corp., "The Vana System," Report, Littleton, Colorado (1983).

Antell, R.E., "The Laser-Scan Fastrak Automatic Digitising System," In: Fifth International Symposium on Computer-Assisted Cartography, August, 1982, 51-64.

Boyle, A.R., "Scan digitization of cartographic data: Graphic Systems Design and Application Group," University of Saskatchewan Electrical Engineering Department, Report, 1979.

Broomall Industries, "AGDS Quick Reference Manual," DMAHTC.

Broomall Industries, "Scanner A, Operating Procedure," In: AGDS Scanner Sub-System User's Manual (DMAHTC).

Callahan, George M., "Digital Cartographic Data Production Using Raster Scanning Techniques," In: Automated Cartography: Auto-Carto Six, Volume 1, October, 1983, 392-399.

Comeau, M.A., "An Investigation of Cartographic Map Scanners," CLDS Technical Bulletin No. 6, September, 1981.

Comeau, M.A., "Functional Components of Cartographic Raster Scanning Systems" CLDS Technical Bulletin No. 7, March, 1983.

Fulford, M.C., "The Fastrak Automatic Digitising System," Pattern Recognition, Vol. 14, No. 1-6, 65-74.

Gibson, Laurie, and Lucas, Dean, "Automated Data Entry System for Digital Cartographic Applications," Technical Papers of the American Society of Photogrammetry (48th Annual Meeting) ACSM-ASP Convention, 1982.

Holecheck, Rosalene M., and Rosenthal, Richard, Mini Raster to Vector Carto Data Validation and Plotting (Project Summary), U.S. Army Engineer Topographic Laboratory, 1983.

Ida, Yoshiaki, "Automatic Digitizer of Topographic Contour Maps" Society Japan, Vol. 35, No. 4, 557-565.

Leberl, Franz W., "Raster Scanning for Operational Digitizing of Graphical Data," Programmetric Engineering and Remote Sensing, Vol. 48, No. 4, April, 1982, 615-627.

Lichtner, W., "Experiments on Automatic Digitizing of Cartographic Maps by Means of Color-Scanner," In: ISPRS Commission Symposium, 1982, Vol. 24-IV, 1982.

Lum, R.M., and Jones, R.P., "Microcomputer-controlled digitizer and image analyzer," Review of Scientific Instruments (Rev. Sci. Instrum.), Vol. 51, No. 7, 954-958.

Nagy, George, "Optical Scanning Digitizers," Computer, May, 1983.

Naylor, Alonzo, D., Trip Report, July, 1983.

Peuquet, Donna, J., and Boyle, A., Raster Scanning, Processing and Plotting of Cartographic Documents, SPAD Systems, Ltd., 1984.

Petrie, G., "Hardware Aspects of Digital Mapping," Photogrammetric Engineering and Remote Sensing, Vol. 47, No. 3, March 1981, 307-320.

Pferd, W. and Stocker, G.C., "Optical Fibers for Scanning Digitizers," The Bell System Technical Journal, Vol. 60, No. 4, April, 1981, 523-534.

Rhind, David, et al., "Towards A National Digital Topographic Data Base: Experiments in Mass Digitizing, Parallel Processing and The Detection of Change," In: Automated Cartography: Auto-Carto Six Vol. 1, October 1983.

Stover, Richard N., "Automated Database Capture for CADD", Proceedings of the Sixth International Symposium on Automated Cartography, 1983, 398-405.

Tafoya, Benny R., "A Unique Cartographic, Scanning, Vectorizing, and Editing System," SPIE, Vol. 155, 1978.

Theis, Richard F., "Raster Data Acquisition and Processing," In: Fifth International Symposium on Computer-Assisted Cartography, August 1982, 667-676.

Vogel, James M., "In Laboratory Investigative Research Intelligent Live Digitizing Study," Interim Report by U.S. Engineer Topographic Laboratories Mapping Development Division, March 1983.

Williams, C.M., "Scanning Digitizers: Automating the Input of Drawings," Computer Aided Design, Vol. II, No. 4, July 1979, 227-230.

MISCELLANEOUS

Clark, Richard A., "Cartographic Raster Processing Programs at USAETL," In: Technical Papers ACSM-ASP Convention, 1980.

Crane, R.K., "Automatic Cell Detection and Tracking," IEEE Transactions Geoscience Electronics, GE-17, 1979.

Fegas, Robin G., "Modularization of Digital Cartographic Data Capture," Proceedings of the Fifth International Symposium on Automated Cartography, 1982, 297-307.

Fegas, Robin G., "Topological Structuring of Raster-Scanned Line Map Data," Technical Papers of the 44th Annual Meeting of the American Congress on Surveying and Mapping (ACSM), 1984, 265-272.

Goodchild, Michael F., and Grandfield, Andrew W., "Optimizing Raster Storage: An Examination of Four Alternatives," In: Automated Cartography: Auto-Carto Six Vol. 1, October 1983, 400-407.

Jenks, Dr. George, "Lines, Computers, and Human Frailties," Annals of the Association of American Geographers, March 1981, Vol. 71, Number 1, 1-10.

Lucas, Dean and Gibson, Laurie, "Introduction to GBT," Report by Interactive Systems Corporation.

Lucas, Dean and Gibson, Laurue, "A Hierarchical Pattern Extraction System for Hexagonally Sampled Images," Report by Interactive Systems Corp.

Gottschalk, H., "The Manipulation of Raster Data," Automation in Cartography, Inter. Cart. Assoc., 1975, 241-253.

Jacobson, J.D. "Geometric Relationships for Retrieval of Geographic Information," IBM Systems Journal, Vol. 7, 1968, 331-341.

Kolassa, David A., "The Evolution of Raster Processing Technology Within the Cartographic Environment," In: Fifth International Symposium on Computer-Assisted Cartography, August 1982, 451-460.

Lybanon, Mathew, Brown, Robert M. and Gronmeyer, Larry K., "Recognition of Handprinted Characters for Automated Cartography: A Progress Report," Proceedings of the Society of Photo-Optical Instrumentation Engineers, Vol. 205, 1980, 165-174.

Miller, Stephen W., "A Raster-Encoded Polygon Data Structure for Land Resource Analysis Applications," In: Fifth International Symposium on Computer-Assisted Cartography, August 1982, 529-536.

Opitz, Bruce K., Becker, Jerry and Cook, Henry, "The Defense Mapping Agency's Pilot Digital Operations," In: Technical Papers of the American Society of Photogrammetry, March 1982, 285-294.

Peuquet, Donna J., "Raster Processing: An Alternative Approach to Automated Cartographic Data Handling," The American Cartographer, Vol. 6, No. 2, 1979, 129-139.

Schneier, Jan S., "Chart Modification and Reproduction with a Raster System," Technical Papers of the American Congress on Surveying and Mapping, 40th Annual Meeting, 1980, 71-77.

Smith, Doyle G., "Raster Data Development in the National Mapping Division U.S. Geological Survey," USGS paper.

RASTER-TO-VECTOR CONVERSION

Adams, T.A. and Smith, K., "The Manipulation of Raster-Based Topographic Data on a Parallel Processor," Proceedings of the IEEE Conference on Pattern Recognition and Image Processing, Las Vegas, 1982, 396-404.

Agrawala, A.K., and Kulkarni, A.V., "A Sequential Approach to the Extraction of Shape Features," Computer Graphics and Image Processing, Vol. 6, 1977, 538-557.

Arcelli, C., and Levialdi, S., "Concavity Extraction by Parallel Processing," IEEE Transactions on Systems, Man, and Cybernetics, October 1971, 394-396.

Boyle, Ray, "Scan Digitizing and Drafting," Technical Papers of ACSM-ASP Convention, 1980.

Clarke, Keith C., "Polygon to Raster Conversion, Error: A Comparative Analysis of Techniques," Seventeenth International Symposium on Remote Sensing of Environment, 1983, 1-10.

Comeau, M.A. "Compression and Compaction of Binary Raster Images," In: Automated Cartography: AutoCarto-Six: Vol. 1, October 1983, 362-371.

Crane, Robert, "Mini Raster-to-Vector Conversion," Report for Engineer Topographic Laboratories, September 1981.

Deutch, E.S., "Thinning Algorithms on Rectangular, Hexagonal, and Triangular Arrays," Communications of the Association for Computing Machinery, Vol. 15, 1972, 827-837.

Goodyear Aerospace Corp., "Associative Array Processing of Raster Scanned Data for Automated Cartography II," Report for U.S. Army ETL, 1977.

Gustafson, Gary B., "Documentation of Mini Raster-to-Vector Conversion Program," Report prepared for U.S. Army ETL, October 1981.

Hsu, S. and Huang X, "Construction of Polygon Maps with Raster-Vector Data Base Conversion Method," Technical Papers, ACSM-ASP Fall Technical Meeting, 1982, 212-220.

Ibid., "Raster-Vector Conversion Methods for Automated Cartography with Applications in Polygon Maps and Feature Analysis," In: Fifth International Symposium on Computer-Assisted Cartography, August 1982, 407-418.

Jackson, H., "AGDS Thinning, Noise Filter Thinning Program Algorithms and Mathematical Development," DMAAC Internal Report, September 1983.

Kreifelts, T., "Skeletonizing and Line Following in Raster-Digitized Line Structures," In: Proceedings of GI/NTG Congress on Digital Image Processing, Munich, Germany, 1977.

Loew, M.H., Pickhotz, R. L., Goldman, L., "Image Processing for Cartographic Applications," Machine Processing of Remotely Sensed Data Symposium, 1981.

Mueller, C.S.M. and Christensen, A.H.J., "TWISK 43 Scanline; A Computer Program for Converting Raster Maps into Line Maps," Technical Report, National Research Institute for Mathematical Sciences.

Nichols, David A., "Conversion of Raster Coded Images to Polygonal Data Structures," Proceedings of the Record VII Symposium, 1981, 508-515.

Peuquet, Donna J., "An Examination of Techniques for Reformatting Digital Cartographic Data: The Raster-to-Vector Process Parts I and II.," U.S. Geological Survey Report, Dec. 1980.

Rome Research Corporation, "Final Technical Report for Automatic Cartographic Feature Identification II," Data Item A004, CLIN 0002, Contract #F30602-C-0063, 1983.

Tafoya, Benny R., "Raster Revisited or Raster-to-Vector and Vector-to-Raster Systems" SPIE, Vol. 222, 1980, 148-155.

Tamura, H., "A Comparison of Line Thinning Algorithms From a Digital Geometry Standpoint," In: Proceedings of the Fourth International Joint Conference on Pattern Recognition, November, 1978.

Wagner, H.L., "Conversion of Raster Coded Images to Polygonal Data Structures," In: Proceedings of Pecora VII Symposium, 1981.

Woetzel G., "A Fast and Economic Scan-to-Line Conversion Algorithm," In: Gesellschaft für Mathematik und Datenverarbeitung MBH Bonn Internal Report, 1977.

COMMERCIAL CONTACT LIST

Edward Getchell
President
ANA Tech Corporation
5321 S. Santa Fe Drive
Littleton, Colorado 80120

Broomall Industries Inc.
700 Abbot Drive
Broomall, Pennsylvania 19008

Elaine V. Barkoulies
Marketing Services Coordinator
Chicago Aerial Survey
2140 Wolf Road
Des Plaines, Illinois 60018

Jeff Schwalm
Marketing Director
Dynamic Graphics
2855 Telegraph Avenue
Suite 405
San Francisco, California

Jack Dangermond
Environmental Systems Research Institute
380 New York Street
Redlands, California 92373

Laurie Gibson
Vice President
Interactive Systems
5500 S. Sycamore
Littleton, Colorado 80210

David Sinton
Intergraph Corporation
One Madison Industrial Park
Huntsville, Alabama 35807

Jeppeson and Sanderson
55 Iverness Drive, East
Englewood, Colorado 80112-5498

Robert D. Pettit
General Manager
IDM-TOWILL
1765 Scott Boulevard
Santa Clara, CA 95050

James W. Wills
Marketing Manager and
Dr. Gunter Desterhelt
Technical Director
Kongsberg North America, Inc.
SysScan Map Information Systems
Service Bureau/Demo Center
9001 Edmonston Road
Suite 402
Greenbelt, Maryland 20770

Dr. Tim Adams
Cartographic Systems Advisor
Laserscan Laboratories Limited
Cambridge Science Park
Milton Road
Cambridge CB44BH
England

Franz W. Leberl
System Software Development
Markhurd Aerial Surveys Inc.
345 Pennsylvania Avenue South
Minneapolis, Minnesota 55426

Steven K. King
Vice President
CAD/CAM Systems
Scitex America Corporation
Eight Oak Park Drive
Bedford, Massachusetts 01730

Samuel Len
Software Director
Spad Systems
P. O. Box 571
Williamsville, New York 14221

Lowell Nerenberg
Sales Engineer
Gerber Scientific Instrument Company
1370 Piccard Drive
Rockville, MD 20850

David J. Baker
Director Applied Technology
Bob Gogineni
Director of Research and Development and Engineering
TELEDYNE GEOTRONICS
725 East Third Street
P.O. Box 480
Long Beach, CA 90801

PROFESSIONAL CONTACT LIST

R. A. Abernathy
Assistant Librarian
University of Durham
University Library
Stockton Road, Durham (0385) 64971
England

George M. Callahan
Research Scientist for Advanced
Cartographic Systems
U. S. Geological Survey
National Mapping Division
Reston, Virginia

Franz W. Leberl
Technical University of Graz
Research Center
A-8010 Graz, Austria

Mr. Jim Malone
Division of National Mapping
WINH 31C
Cameron Offices
P. O. Box 13
Bel Connen, Australia

Mr. Neil S. Smith
Ordnance Survey
Maybush UK
Southampton 509 4OH

Professor Pinhas Yoeli
Tel Aviv University
Department of Geography
Tel-Aviv 69978, Israel

Appendix D

SYSTEM PROFILE QUESTIONNAIRE

COMPANY ______________________________

OVERALL SYSTEM CHARACTERIZATION

Official System Name ______________________________

Hardware Components:

- CPU ______________________________
- Scanner ______________________________
- CRT's ______________________________
- Data Storage ______________________________
- Tape Drives ______________________________
- Disk Drives ______________________________
- Manual Digitizing Units ______________________________

Software Components:

- System Packages ______________________________
- Programmable Languages/ Capabilities ______________________________

Installation Requirements:

- Floor Space ______________________________
- Power Supply ______________________________
- Water Supply ______________________________
- Environment ______________________________
- Computer ______________________________

Miscellaneous Information:

- Aquisition Cost ______________________________
- Maintenance Cost (contract) ______________________________
- Support Facilities ______________________________
- Research and Development ______________________________
- Systems Sold/In Production ______________________________
- Training ______________________________
- System Upgrade/Expansion ______________________________

Comments: ______________________________

SYSTEM: ______________________

PREPARATION

Manuscript Input Specifications

Materials:

Paper ____

Mylar ____

Scribecote ____

Film Positive ____

Film Negative ____

Miscellaneous specify - ______________________

Image Type:

Pencil ("lead") ____

Color Pencil ____

Black Ink (wet) ____

4-Color Ballpoint ____

Printed Color Line ____

Printed Color Screen ____

Miscellaneous specify - ______________________

Misc. Specifications:

Max. # Colors (1 overly) ________

Min. Lineweight ________

Max. Lineweight ________

Min. Line Separation ________

Max. Sheet Format ____" x ____" / ____mm x ____mm

Max. Image Format ____" x ____" / ____mm x ____mm

Acceptable Generic Data:

Non-intersecting Lines ______________________

Closed Polygons ______________________

Networks ______________________

Point Symbols ______________________

Symbolization simple - ______________________

moderate - ______________________

complex - ______________________

Annotation (Labels) ______________________

Comments: ______________________

SYSTEM: ____________________

DIGITIZATION

System Specifications/Characteristics

Digitizing Method:

Raster Scanning — Electro-Optical ____ Laser ____ Solid State ____

Auto-Line Following — Laser ____ Other ____________

Manual — Standard ____ Line Follow ____ Error Crct. ____

Image Holder Specifications:

Drum ____

Flatbed ____

Fiducial Registration ____

Tab Registration ____

Other Registration — specify - ____________

Sheet Format (max.) — " x " / mm x mm

Image Format (max.) — " x " / mm x mm

Photometric Characteristics:

Color Recognition ____________

Color Storage (1 pass) ____________

Color Calibration ____________

Grey Level Recognition ____________

Grey Level Storage (1 pass) ____________

Grey Level Calibration ____________

Accuracy Specifications:

Geometric Linearity ____________

Stability (Repeatability) — short term- ____________ long term- ____________

Point Accuracy ____________

Point Repeatability ____________

Resolution ____________ Aperture/ ____________

Scan Lines per inch ____________ Spot Size

Productivity Specifications:

Throughput Rate ____________

Speed (Traverse Rate in/sec) ____________

Line-Following Rate ____________

Ease of Use — Simple - ____ Moderate - ____ Complex - ____

Comments: ____________________

____

____

SYSTEM: ______________________________

RASTER-TO-VECTOR CONVERSION

Raw Raster Data Review Edit

Data Storage/Compaction:

Total Row/Column / Scan Line ____

Run Length Encoded ____

Entry/Exit Points ____

Other specify - ______________________________

Raw Data Review:

Computer Graphics Review Only ____

Hard Copy Raster Review Plot ____

Error Detection/Correction:

Error Types	Interactive	Computer-Assisted	Automated
Gaps	____	____	____
Spikes	____	____	____
Stubs	____	____	____
Sticks	____	____	____
Streaks	____	____	____
Snow	____	____	____
Coalescence	____	____	____
Un-straight Lines	____	____	____
Un-square Corners	____	____	____
Missing Data	____	____	____
Geometric Image Distortion	____	____	____

Misc. Specifications:

Ease of Use simple - ___ moderate - ___ complex - ___

Max. Addressable CRT Image ______ pixels ______

Ergonomic Provisions ______________________________

User Friendly Commands ______________________________

Comments: ______________________________

____

____

Raster-To-Vector Conversion

Algorithms:

Skeletonization/ Line Thinning	Ballooning ____ Peeling ____ Medial Axis ____
	Other ____
Line Extraction	Line Following ____ Scan Line ____
	Other ____
Topology Reconstruction/ Vector Data Structure	Link-Node ____ Chain-Enclosed Polygon ____
	SLF ____ SIF ____ Gerber ____ Applicon ____
	CalComp ____ Versatec ____ Other ____
	Commercial Vector Format Name ____

Speed:

Lineal Inches/Minute	Non-intersecting Lines ____ Polygons ____
	Networks ____ Points ____ Average ____

Advanced Facilities:

Automatic Feature Tagging	specify - ____
Character Recognition	specify - ____

Comments: ____

Raw Vector Data Review Edit

Error Detection/Correction:

Error Types	Interactive	Computer-Assisted	Automated
Gaps	____	____	____
Spikes	____	____	____
Stubs	____	____	____
Sticks	____	____	____
Streaks	____	____	____
Slivers	____	____	____
Coincident Points	____	____	____
Coincident Lines	____	____	____
Snow	____	____	____
Unsquare Corners	____	____	____
Un-straight Lines	____	____	____
Displaced Nodes	____	____	____
Skewed Intersections	____	____	____

Wandering "Cenerlines" _____ _____ _____

Topological/Spatial _____ _____ _____

Miscellaneous Specifications:

Max. Addressable CRT Image __________ square map inches

Track Ball _____

Light Pen _____

Cursor/Mouse _____

Color CRT _____

Ease of Use simple - _____ moderate - _____ complex - _____

User Friendly Commands ____________________

Ergonomic Provisions ____________________

Comments: ______________________________

SYSTEM: ____________________

TAGGING

Analog Feature Attribute/
Vector Data Merge

Tagging/Elevating:

Light Pen	____
Cursor	____
Track Ball	____

Data	Interactive	Computer-Assisted	Automated
Contours	____	____	____
Points	____	____	____
Polygons	____	____	____
Networks	____	____	____
Links	____	____	____

Miscellaneous Specifications:

Max. Addressable CRT Image	____________________
Ease of Use	simple - ____ moderate - ____ complex - ____
User Friendly Commands	____________________
Ergonomic Provisions	____________________

Comments: ____________________

____

____

____

SYSTEM: ______________________________

SPATIAL CODING

Universal Referencing

Software for converting machine coordinates to a universal reference system (e.g. UTM, Latitude/Longitude, Alternative Planar Map Projection Systems - Albers Equal Area, Lambert Conic Conformal etc...)

____

____

____

Topological Encoding

Software for developing a truly topological data structure (e.g. Census-DIME, USGS-DLG3, etc...)

____

____

____

Data Structuring

Software for writing a "standard", user-friendly, or DMA-Standard data format (e.g. DMA - Standard Lineal Format) other than strictly proprietary

____

____

____

Comments: ______________________________

____

____

Appendix E

REVIEWS OF GOVERNMENT SPONSORED RASTER-TO-VECTOR CONVERSION RESEARCH AND DEVELOPMENT

Associative Array Processing of Raster Scanned Data for Automated Cartography*

A raster-to-vector conversion capability was developed as a part of this comprehensive raster processing R&D program. The software operated on an associative array processor and took advantage of its unique architecture. The results were demonstrated by processing map sheets through a raster scanner followed by a set of line weight separation, thinning, thickening, raster-to-vector, and automatic editing operations followed by the output of the features on a Gerber plotter. Processing speeds for the raster-to-vector operation were reported to be about four times faster on the associative array processor than on a conventional computer and to be dependent on the resolution of the raster data rather than on the density of the lineal features. The only problem discussed in the report had to do with distortions produced at line junctions.

*"Associative Array Processing of Raster Scanned Data for Automated Cartography", Report No. ETL 0046, Goodyear Aerospace Corporation (March 1976), AD A022 753.

Raster-to-Lineal Conversion Analysis*

In this project a raster-to-vector algorithm was implemented on a HIS 635 and UNIVAC 1108. The process consisted of a skeletonization operation that included the identification and classification of junction points followed by a vectorization operation and then plotting. The vectorization operation is performed in two phases. In the first, raster skeleton strips are processed to form lineal segments for the entire map sheet. This is followed by the connection of the segments to form complete features. Several factors were cited that affect processing rates:

(1) Higher feature densities result in higher rates in terms of lineal feature length per unit of time.

(2) Thicker lines take longer than thin lines due to the thinning operation.

The algorithm apparently produced line breaks and noise features since recommendations included the development of an automatic editing capability.

*"Raster-to-Lineal Conversion Analysis", Report No. RADC-TR-77-421, PRC Information Sciences Company (December 1977).

Mini Raster-to-Vector Conversion*

This raster-to-vector algorithm operates directly on the run length encoded output of a raster scanner. The run length encoding of only two successive scan lines is kept in memory at one time. The process results in the flagging of the unattached ends of spurs, the nodes of intersecting lines, and open lines unattached to the boundary of the map. This flagging is claimed to greatly facilitate the automation of the editing of vectors produced from contour data.

Timing data is presented for both the Data General Eclipse 250 and the DEC PDP 11/60 computers. A number of output examples are shown that illustrate the types of errors and anomalies that were generated from the test data.

*"Mini Raster-to-Vector Conversion", Report No. ETL-0269, Environmental Resarch and Technology, Inc. (September 1981).